The two-dimensional Coulomb gas, vortex unbinding, and superfluid-superconducting films

Petter Minnhagen

Department of Theoretical Physics, Umeå University, S-901 87 Umeå, Sweden

The article reviews the two-dimensional Coulomb gas model and its connection to vortex fluctuations for a two-dimensional superfluid. The neutral and non-neutral versions of the Coulomb gas are discussed and the relation to an equivalent sine-Gordon field theory is given. The charge-unbinding picture is used to elucidate some essential properties of the Coulomb gas. Derivations of approximate renormalization equations are sketched and the phase transition for the neutral two-dimensional Coulomb gas is described. The Kosterlitz renormalization-group equations are reviewed in some detail. The vortex-Coulomb gas charge analogy is carefully explained. The connections between experiments for ^{4}He films and superconducting films and the neutral and non-neutral versions of the Coulomb gas are outlined using concepts like the universal jump and the Coulomb gas scaling relations. The properties of a dynamical version of the Coulomb gas, corresponding to vortex dynamics, are discussed and related to experiments. An analogy with Maxwell's equations in two dimensions is also given.

CONTENTS

GUIDE TO SOME COULOMB GAS SYMBOLS

s $(s=\pm 1)$	single-particle charge
r_0	single-particle extension
$f_{r_0}(r)$	single-particle charge distribution
$E_{c\binom{+}{-}}\Delta E$	creation energy excluding the electrostatic self-energy for a positive (negative) particle
$U(r)$	electrostatic particle interaction
$\frac{1}{2}U(0)$	electrostatic self-energy
$\mathbf{D}$	external electric field
$\mu_{\binom{+}{-}}=\mu_{\binom{-}{+}}\Delta E$	chemical potential for a positive (negative) particle

$\mu=-[\frac{1}{2}U(0)+E_c]$

Copyright ©1987 The American Physical Society

2μ	chemical potential for a neutral pair with large separation
$2\Delta E$	difference in chemical potential between a negative and a positive particle
V	Coulomb potential for a point charge
V_{λ_c} $(V=V_{\lambda_c=\infty})$	Coulomb potential including a background screening
V_L	linearly screened Coulomb potential
λ	screening length (defined through V_L)
λ_c	screening length due to background charge (serves as an infrared cutoff)
Z	grand partition function
ζ	phase-space division
$N_{+}_{(-)}$	number of positive (negative) particles in a configuration
$N=N_+ + N_-$	total number of particles in a configuration
H_N	configuration energy
T	temperature
Ω	volume
$z_{+}_{(-)}$ $(z_{+}_{(-)}=e^{\mu_{+}_{(-)}/T})$	fugacity for a positive (negative) particle
$z=e^{-E_c/T}$	
z^2	$\simeq$ fugacity for a neutral pair with small separation
$n_F^{+}_{(-)}$	density of free positive (negative) particles
$\langle \Delta n \rangle$	charge density. The charge density for a configuration of N particles is given by $\Delta n(r)=\sum_i^N s_i f_{r_0}(\lvert \mathbf{r}-\mathbf{r}_i \rvert)$.
$\tilde{\varepsilon}$	dielectric constant describing the polarization due to bound pairs
$\varepsilon(r)$	linear-response dielectric function
ε_∞ $[\varepsilon_\infty=\hat{\varepsilon}(k=0)]$	dielectric constant (a caret denotes Fourier transform)
ε_r	length-dependent dielectric constant
F	free-energy density
F_B	free-energy density due to bound pairs
F_F	free-energy density due to free particles

I. INTRODUCTION

This article is an attempt to describe the physics of the Kosterlitz-Thouless transition and its connection to some physical realizations for superfluid films and superconducting films. It is not intended as a review in the sense that every possible angle is covered. Rather it is a review based on a single perspective, the Coulomb gas perspective.

The physics of the two-dimensional Coulomb gas is especially easy to grasp and visualize. By expressing as much as possible in terms of this model one gains an intuitive understanding that goes quite far. Hopefully, a review based on this perspective may be of some use to a more general audience.

To some extent it may also help to bridge a gap between the extremes of the esoteric theorist on the one side and the practical experimentalist on the other; this gap may, perhaps, sometimes take the form of two quite different persons having trouble communicating. Two such persons will, however, undoubtedly agree that the present article treats aspects of their respective special interests quite inadequately, if indeed at all. Nevertheless, my hope is that such shortcomings to some extent are compensated for by the intuitive insight gained by the Coulomb gas perspective.

The subject derives from the idea that the quasi-long-range order for two-dimensional XY-like systems is destroyed by vortex unbinding (Berezinskii, 1971; Kosterlitz and Thouless, 1972, 1973). This will be referred to as the Kosterlitz-Thouless (KT) transition following what seems to be the most common usage. The renormalization-group equations for the Kosterlitz-Thouless transition were constructed by Kosterlitz (1974). This led to the prediction of a universal jump of the superfluid density for superfluid films (Nelson and Kosterlitz, 1977). This prediction was tested against experiments on ^{4}He films and good agreement was found (Bishop and Reppy, 1978). A dynamic theory was developed (Ambegaokar *et al.*, 1978, 1980). It was realized that these ideas should also apply to superconducting films (Beasley *et al.*, 1979; Doniach and Huberman, 1979; Halperin and Nelson, 1979; Turkevich, 1979).

Outlined in this way the subject has been covered from various angles and to various degrees of completeness in several review articles, e.g., those of Kosterlitz and Thouless (1978), Halperin (1979), Nelson (1980, 1983), Hebard and Fiory (1982), Minnhagen (1984a), and Mooij (1983, 1984).

A considerable amount of the work done on the subject so far consists of attempts to tie various experimental observations to the predictions obtained from Kosterlitz renormalization-group equations or to the dynamical theory of Ambegaokar *et al.* (1978, 1980), which is based on these renormalization-group equations.

The present article puts more emphasis on the properties of the two-dimensional Coulomb gas *per se* and less on some of the more specific results obtained from the Kosterlitz renormalization-group equations. One reason for this is that the critical region to which the renormalization-group equations strictly apply may be very narrow and, possibly, to a large extent inaccessible to experiments. From this point of view, more general and

less detailed properties of the two-dimensional Coulomb gas may have a larger validity range in relation to the experiments and may, because of this, provide a firmer link to experiments.

The subject is in many respects still quite open, especially as regards the precise connection between theory and experiments. This article focuses on what may be judged to be the core of the present understanding of the subject.

The general outline of the article is as follows: Sec. II describes the two-dimensional Coulomb gas model, its relation to a sine-Gordon field theory, and expansion approaches. The renormalization-group equations, together with their limitations and possible extensions, are discussed in Sec. III. The connection to vortex fluctuations for two-dimensional superfluids is given in Sec. IV, while Sec. V contains examples of experiments on ^{4}He films and superconducting films reflecting "static" Coulomb gas properties. Dynamical aspects are discussed in Sec. VI, and examples of experiments reflecting "dynamical" Coulomb gas properties are given. Section VII contains various comments. The article is concluded by Sec. VIII, where some final remarks are given.

II. THE TWO-DIMENSIONAL COULOMB GAS

The two-dimensional Coulomb gas model is a statistical mechanics model defined in terms of a grand partition function. The particles in the model are Coulomb gas particles with equal or opposite charge, interacting through a Coulomb interaction that, in two dimensions, turns out to be logarithmic. The two-dimensional Coulomb gas model undergoes a charge-unbinding transition in a narrow temperature region, in which particles bound together in neutral pairs unbind. The subject of the present article will be centered on this charge-unbinding transition. In a certain limiting case of the two-dimensional Coulomb gas model, all particles are bound together in pairs for low enough temperatures, and the charge-unbinding transition starts at a critical temperature at which the model undergoes a thermodynamic phase transition, the Kosterlitz-Thouless transition.

The definition of the two-dimensional Coulomb gas model used in this article is given in Sec. II.A, and an important distinction between two limiting cases is made, i.e., the neutral Coulomb gas and the non-neutral Coulomb gas. Section II.B contains heuristic derivations of the main properties of the neutral and non-neutral Coulomb gas. A sine-Gordon field theory formulation of the two-dimensional Coulomb gas is given in Sec. II.C. Expansion approaches for obtaining the properties of the model in a more stringent way are described in Sec. II.D, while Sec. II.E comments on some fermion analogies.

A. Definitions

A Coulomb gas consists of particles with positive or negative charge of equal magnitude. The charge is denoted by s, and units are chosen such that $s=\pm 1$. The charges interact via the Coulomb interaction, which is defined by Poisson's equation,

$$\nabla^2 V(r) = -2\pi\delta(\mathbf{r}) , \tag{2.1}$$

where $V(r)$ depends on the space dimension (iD, $i=1,2,3$),

$$V(r) \sim \begin{cases} r, & 1\text{D} , \\ \ln(r), & 2\text{D}, \\ 1/r, & 3\text{D} . \end{cases} \tag{2.2}$$

Consequently the interaction for the two-dimensional Coulomb gas depends logarithmically on distance. One may also note from Eq. (2.2) that the dimension of charge in two and three dimensions differs by a factor (length)$^{1/2}$, since in both cases $s^2V(r)$ has the dimension of energy, i.e., s(two-dimensions)$=s$(three-dimensions)/(length)$^{1/2}$. Thus in two dimensions s^2 has the dimension of energy.

1. The Coulomb gas model

We proceed by defining precisely the two-dimensional Coulomb gas model to be used throughout the present article. It should be noted that some properties do not depend on the precise details, whereas others do, which will be made clearer as we go along.

In the model a charge is associated with a spatial charge distribution, $f_{r_0}(r)$, identical for all charges. The linear extension of a charge is given by r_0, and $f_{r_0}(r)$ fulfills the single charge normalization condition $\int d^2r\, f_{r_0}(r)=1$. The limit $r_0=0$ is the point-charge limit, $f_{r_0=0}(r)=\delta(\mathbf{r})$. In the field theory formulation of the two-dimensional Coulomb gas (see Sec. II.C) r_0 plays the role of an ultraviolet cutoff (Minnhagen *et al.*, 1978; Samuel, 1978). Similarly, it is expedient to include an infrared cutoff in the definition of the particle interaction. This will be implemented by changing the definition of the interaction given by Eq. (2.1) into $V_{\lambda_c}(r)$, with

$$(\nabla^2-\lambda_c^{-2})V_{\lambda_c}(r) = -2\pi\delta(\mathbf{r}) . \tag{2.3}$$

The length λ_c may be interpreted as the Debye screening length arising from a neutral background of infinitesimal charges (Minnhagen, 1981a). In some applications for superfluid-superconducting films, to be discussed later, the length λ_c will play an essential physical role.

With these definitions the interaction energy between two Coulomb gas charges i and j at positions $\mathbf{r}_i$ and $\mathbf{r}_j$ and with charges s_i and s_j, respectively, is given by $s_is_jU(|\mathbf{r}_i-\mathbf{r}_j|)$, with the interaction U defined by

$$U(r) = \int d^2r'd^2r''f_{r_0}(|\mathbf{r}-\mathbf{r}'|)V_{\lambda_c}(|\mathbf{r}'-\mathbf{r}''|)f_{r_0}(r'') . \tag{2.4}$$

The interaction $U(r)$ has the following limiting forms:

$$U(0)\simeq\ln(\lambda_c/r_0) \quad \text{for } \lambda_c/r_0 \gg 1 , \tag{2.5a}$$

$$U(r)\simeq -\ln(r/\lambda_c) \quad \text{for } r_0 \ll r \ll \lambda_c , \tag{2.5b}$$

$$U(r)\sim \frac{1}{\sqrt{r}} e^{-r/\lambda_c} \quad \text{for } \lambda_c \ll r , \tag{2.5c}$$

where $\simeq$ implies the leading term, whereas $\sim$ implies the dominating r dependence. The electrostatic self-energy of a Coulomb gas particle is defined by $\frac{1}{4}\int d^2r\, \mathbf{E}^2(r)$ where $\mathbf{E}$ is the electric field caused by a single Coulomb gas charge. This self-energy can be expressed as $\frac{1}{2}U(0)$, using standard electrostatics. The single-charge self-energy diverges in the limit $\lambda_c \to \infty$ [compare Eq. (2.5a)].

The electrostatic energy associated with a configuration of N Coulomb gas charges is given by

$$\tfrac{1}{2}\sum_{ij} s_i s_j U(r_{ij}) ,$$

where $r_{ij}=|\mathbf{r}_i-\mathbf{r}_j|$ and the indices i and j run over all the Coulomb gas charges.

In the physical realizations of the two-dimensional Coulomb gas to be discussed later, the Coulomb gas particles are created as thermal excitations. The creation of such an excitation will, in addition to the electrostatic contribution, involve a nonelectrostatic contribution, which may be associated with the local formation of the single-charge distribution on the scale of r_0. This additional contribution we shall take to be a constant denoted by E_c. From the point of view of the Coulomb gas model, this means that the chemical potential of a Coulomb gas charge, μ, is given by

$$\mu=-[\tfrac{1}{2}U(0)+E_c] . \tag{2.6}$$

A further generalization is the case in which the energies needed to create a positive and a negative Coulomb gas charge are unequal. This means that the chemical potentials of positive, μ_+, and negative, μ_-, charges are unequal and are given by

$$\mu_{\substack{+\\(-)}} = -[\tfrac{1}{2}U(0)+E_{c(\substack{+\\-})}\Delta E] , \tag{2.7}$$

where $2\Delta E$ is the difference in creation energy for a positive and negative Coulomb gas charge.

The discussion in this article will be based on a Coulomb gas model for which the total energy H_N associated with a configuration of $N_{\substack{+\\(-)}}$ positive (negative) charges, $N=N_+ +N_-$, is given by

$$H_N=\tfrac{1}{2}\sum_{i\neq j} s_i s_j U(r_{ij})-N_+\mu_+-N_-\mu_- . \tag{2.8}$$

Central to the subject under consideration are the thermodynamic properties of this two-dimensional Coulomb gas in the limit of large λ_c/r_0. These are contained in the partition function Z defined by

$$Z=\sum_{N_+,N_-=0}\frac{1}{N_+!}\frac{1}{N_-!}\prod_i \frac{\int d^2r_i}{\zeta} e^{-H_N/T} . \tag{2.9}$$

Here $e^{-H_N/T}$ is the Boltzmann factor for a configuration of $N=N_+ +N_-$ charges containing $N_{\substack{+\\(-)}}$ positive (negative) charges. T is the temperature in units of s^2 [s^2 has the dimension of energy; compare the paragraph below Eq. (2.2)]. The sum is over configurations with all possible values of N_+ and N_-. The index i is assigned to the individual charges in a configuration specified by N_+ and N_-, $\mathbf{r}_i$ is the position of these charges, and, finally, ζ is a constant phase-space division. The size of ζ in the Coulomb gas model is given by $\zeta=\tilde{c}r_0^2$, where r_0 is the linear extension of a charge and $\tilde{c}$ is a constant. Equation (2.9) completes the definition of the Coulomb gas model.

The thermodynamics of this model may be described in terms of three independent variables, the temperature T and the fugacities for positive, z_+, and negative, z_-, charges (defined by $z_\pm=e^{\mu_\pm/T}$). In terms of these variables the partition function becomes

$$Z=\sum_{N_+,N_-=0}\frac{z_+^{N_+}}{N_+!}\frac{z_-^{N_-}}{N_-!} \times\prod_i\frac{\int d^2r_i}{\zeta}\exp\left[-\sum_{i<j} s_i s_j U(r_{ij})/T\right] , \tag{2.10}$$

where the sum in the exponent runs over all pairs of charges within a configuration specified by N_+ and N_-. Z is to lowest order in $z_{\substack{+\\(-)}}$ given by

$$Z=1+\frac{\Omega}{\zeta}z_+ +\frac{\Omega}{\zeta}z_- +\cdots , \tag{2.11}$$

where Ω is the (2D) volume. The thermal average number of positive (negative) charges, $\langle N_{\substack{+\\(-)}}\rangle$, is related to Z by

$$\langle N_{\substack{+\\(-)}}\rangle = z_{\substack{+\\(-)}}\frac{\partial}{\partial z_{\substack{+\\(-)}}}\ln(Z) , \tag{2.12}$$

which to lowest order in $z_{\substack{+\\(-)}}$ becomes

$$\langle N_{\substack{+\\(-)}}\rangle = \frac{\Omega}{\zeta}z_{\substack{+\\(-)}} . \tag{2.13}$$

Equation (2.13) has the following intepretation: $z_{\substack{+\\(-)}}$ is the probability for the process of creating one positive (negative) charge in a particular phase-space division, and Ω/ζ is the total number of such divisions. This clarifies the physical meaning of $z_{\substack{+\\(-)}}$ and ζ. The higher-order terms in $z_{\substack{+\\(-)}}$ for $N_{\substack{+\\(-)}}$ give the corresponding probabilities for the processes of creating more than one charge at a time.

2. Neutral and non-neutral Coulomb gas

The configuration energy H_N can be expressed as

$$H_N=\tfrac{1}{2}\sum_{i\neq j}s_is_j[U(r_{ij})-U(0)] +\tfrac{1}{2}(N_+-N_-)^2U(0)+NE_c+(N_+-N_-)\Delta E \tag{2.14}$$

by use of the identity

$$\tfrac{1}{2}\sum_{i\neq j}s_is_j=\tfrac{1}{2}(N_+-N_-)^2-\tfrac{1}{2}N \tag{2.15}$$

[compare Eqs. (2.7) and (2.8)]. Since $U(r)-U(0)\simeq-\ln(r/r_0)$ for $r_0\ll r\ll\lambda_c$ [compare Eqs. (2.5)], whereas $U(0)\simeq\ln(\lambda_c/r)$, it follows that non-neutral configurations (i.e., configurations with $N_+\neq N_-$) will correspond to a huge energy in the large λ_c/r_0 limit, due to the term proportional to $U(0)$ in Eq. (2.14) as compared to the neutral ones (i.e., the ones with $N_+=N_-$). Consequently, in the limit $\lambda_c/r_0\to\infty$ the thermodynamic properties will be determined by the neutral configurations. This means that the Coulomb gas model reduces to the neutral Coulomb gas, for which the energy of a configuration is given by

$$H_N\simeq-\tfrac{1}{2}\sum_{i\neq j}s_is_j\ln(r_{ij}/r_0)-NT\ln(z)\ , \tag{2.16}$$

where $z=e^{-E_c/T}$ [compare Eq. (2.14)] and only neutral configurations are allowed.

Expanding Z [see Eq. (2.9)] in z gives the lowest-order term,

$$Z=1+\frac{\Omega z^2}{\zeta}\int\frac{dr}{\zeta}2\pi re^{[U(r)-U(0)]/T}\ . \tag{2.17}$$

The leading behavior of $e^{[U(r)-U(0)]/T}$ for large r is $(1/r)^{1/T}$ [compare Eq. (2.5a)]. Thus the expansion in Eq. (2.17) gives the leading small-z dependence for $T<\frac{1}{2}$. In analogy with Eq. (2.13) this small-z contribution is related to the number of bound pairs with pair separation in the interval $[r,r+dr]$, $\langle N_b(r)\rangle$, i.e.,

$$\langle N_b(r)\rangle=\frac{\Omega z^2}{\zeta^2}2\pi r\,dr\,e^{[U(r)-U(0)]/T}\ \text{ for } T<\tfrac{1}{2}\ . \tag{2.18}$$

An alternate way of defining the neutral Coulomb gas, which is common in the literature, is to take H_N given by Eq. (2.16) as the definition of the configuration energy and impose the restriction that two charges cannot be closer than r_0. This would correspond to charges that were impenetrable disks. The phase-space division is usually taken to be r_0^2.

The key feature common to all definitions of the neutral two-dimensional Coulomb gas is a logarithmic interaction that is cut off at some small interparticle distance. This is enough to ensure identical critical properties at the Kosterlitz-Thouless transition. The critical properties are *universal* in this sense. This will be one use of the term "universal" in the present article. The noncritical properties will then be nonuniversal in the sense that they will depend on the precise definitions, e.g., on precisely how the logarithmic interaction is cut off at small distances.

The word universal is sometimes also connected with the noncritical properties in another sense, i.e., different physical systems may be well described by the very same two-dimensional Coulomb gas model. This will be discussed further in connection with superfluid-superconducting films. A model of interest in this context is the Ginzburg-Landau Coulomb gas model, which corresponds to a specific charge distribution $f_{r_0}(r)$, E_c, and ζ (see Sec. IV.D).

The non-neutral Coulomb gas is obtained in the large-λ_c/r_0 limit provided that $\mu_+-\mu_-\sim\lambda_c^2$ (Minnhagen, 1980, 1981a). The properties of this model will be of interest, for example, in connection with superconducting films (compare Sec. IV.B.1 and Table III). The non-neutral Coulomb gas model does not undergo a Kosterlitz-Thouless transition in the thermodynamic phase transition sense and consequently does not lead to any universal critical properties caused by charge unbinding. Nevertheless, charge unbinding is a dominant feature of the model.

B. Heuristic derivation of basic properties

In this section the basic properties of the neutral and non-neutral two-dimensional Coulomb gas are described and motivated through heuristic reasoning. Two Coulomb gas quantities will be of particular interest throughout the article. These are the Debye screening length λ and the dielectric constant $\tilde{\varepsilon}$ (precise definitions of these quantities are given at the beginning of Sec. II.B.1). The dominant temperature dependencies caused by charge unbinding will be discussed.

Section II.B.1 describes the charge-unbinding transition for the neutral two-dimensional Coulomb gas; the most important results are summarized in Table I. Section II.B.2 generalizes the result for the Debye screening length to the case of a non-neutral two-dimensional Coulomb gas. Finally, Sec. II.B.3 gives a heuristic argument for the dependence of the screening length on an external electric field for temperatures below the charge-unbinding transition in the case of a neutral two-dimensional Coulomb gas.

1. Neutral Coulomb gas

We begin by giving a characterization of the Kosterlitz-Thouless transition together with definitions of the screening length λ and the dielectric constant $\tilde{\varepsilon}$. The charge-unbinding picture is introduced. Then we shall describe and give motivation for the basic features of the charge-unbinding transition. Section II.B.1.a contains a heuristical motivation of a basic equation for the screening length λ. Section II.B.1.b describes the critical

behaviors of the dielectric constant $\tilde{\varepsilon}$ and the screening length λ. The Coulomb gas universal jump is explained in Sec. II.B.1.c, and, finally, Sec. II.B.1.d describes and motivates the leading behavior of the free-energy density in the limit of small fugacity z.

The neutral two-dimensional Coulomb gas undergoes a Kosterlitz-Thouless transition in the $\lambda_c \to \infty$ limit for fixed r_0 (Berezinskii, 1971; Kosterlitz and Thouless, 1973). One way of describing this transition is with the aid of the linearly screened potential $V_L(r)$. This is the potential outside an infinitesimal test charge inserted into the Coulomb gas. The screening and polarization of this potential caused by the Coulomb gas charges reflect the properties of the Coulomb gas *per se*. To be precise we shall define $V_L(r)$ as the potential per unit test charge outside an infinitesimal test charge with a Coulomb gas single-particle charge distribution $f_{r_0}(r)$. The quantity $V_L(r)$ will be a key quantity in the present article.

The energy needed to separate two infinitesimal test charges with charges $+\delta t$ and $-\delta t$, respectively, by a distance r is related to $V_L(r)$ by

$$(\delta t)^2 \int d^2r' f_{r_0}(r')[V_L(r') - V_L(|\mathbf{r}-\mathbf{r}'|)] .$$

In terms of this separation energy the Kosterlitz-Thouless transition may be characterized in the following way: For every fixed fugacity z there exists a critical temperature $T_c(z)$ such that the separation energy between the test charges falls off exponentially with separation for $T > T_c(z)$, whereas it grows logarithmically with separation for $T < T_c(z)$.

This change in behavior at T_c may be given an intuitive physical interpretation (Kosterlitz and Thouless, 1973): Below T_c *all* Coulomb gas charges are bound in pairs, whereas above T_c not all charges are bound in pairs. In other words, as the temperature is raised above T_c, neutral pairs of particles unbind. As a result there are free Coulomb gas charges present above T_c and these free charges screen the test-charge interaction, causing an exponential falloff at large separations. The transition is quite commonly called the Kosterlitz-Thouless charge-unbinding transition, in reference to this intuitive physical interpretation.

It is instructive to establish this charge-unbinding picture at the level of a Poisson-Boltzmann description. The potential between the test charges may, within such a description, be expressed as

$$\nabla^2 \delta t V_L = -\frac{2\pi\delta t}{\tilde{\varepsilon}} f_{r_0} - \frac{2\pi n_F^+}{\tilde{\varepsilon}} e^{-\delta t V_L/T} + \frac{2\pi n_F}{\tilde{\varepsilon}} e^{\delta t V_L/T} . \tag{2.19}$$

Here $n_F^{(+)}$ is the density of positive (negative) Coulomb gas charges and $\tilde{\varepsilon}$ is the dielectric constant due to the polarization caused by the bound pairs of Coulomb gas charges. Now, since δt is infinitesimal, Eq. (2.19) reduces to

$$\nabla^2 \delta t V_L = -\frac{2\pi\delta t}{\tilde{\varepsilon}} f_{r_0} - \frac{2\pi\delta t n_F}{\tilde{\varepsilon} T} V_L , \tag{2.20}$$

where $n_F = n_F^+ + n_F^-$ is the total density of free charges. In terms of Fourier transforms this becomes (a caret above a function denotes the Fourier transform)

$$\hat{V}_L(k) = \frac{1}{\tilde{\varepsilon}} \frac{2\pi}{k^2 + \lambda^{-2}} \hat{f}_{r_0}(k) , \tag{2.21a}$$

where

$$\lambda^{-2} = \frac{2\pi n_F}{\tilde{\varepsilon} T} . \tag{2.21b}$$

It follows [note that $f_{r_0}(k=0)=1$ by definition] that

$$V_L \sim \begin{cases} \dfrac{1}{\sqrt{r}} e^{-r/\lambda}, & \lambda \neq \infty , \quad (2.22a) \\ -(1/\tilde{\varepsilon}) \ln(r), & \lambda = \infty . \quad (2.22b) \end{cases}$$

Within the charge-unbinding picture, the screening length λ is related to the presence of free Coulomb gas charges, whereas the dielectric constant $\tilde{\varepsilon}$ is related to the polarization due to bound pairs. The leading small-k dependence for V_L is always of the form given by Eq. (2.21a) and, consequently, the $k \to 0$ limit of V_L provides a precise definition of λ and $\tilde{\varepsilon}$. Similarly, Eq. (2.21b) will be taken as the precise definition of the free-particle density n_F. It may be noted that, since λ and $\tilde{\varepsilon}$ can be defined by the small-k limit of V_L, the characterization of the Kosterlitz-Thouless transition in terms of λ and $\tilde{\varepsilon}$ does not presume the charge-unbinding interpretation.

A heuristic argument by Kosterlitz and Thouless (1973) as to why the neutral two-dimensional Coulomb gas would undergo a phase transition is based on the balance between the entropy and the electrostatic self-energy associated with introducing a single Coulomb gas particle. Above a certain temperature the increase in entropy wins out over the electrostatic self-energy, whereas below the reverse is true. The argument goes as follows: Consider the free-energy density F_1, for a single Coulomb gas particle in a medium of bound pairs, in the limit $\lambda_c \to \infty$ for fixed volume Ω. It is given by

$$\begin{aligned} F_1 &= \frac{1}{2\tilde{\varepsilon}} U(0) - TS \\ &\simeq \frac{1}{2\tilde{\varepsilon}} \ln(R/r_0) - T \ln(R^2/r_0^2) \\ &\simeq \left[\frac{1}{4\tilde{\varepsilon} T} - 1 \right] \ln(R^2/r_0^2) , \end{aligned} \tag{2.23}$$

where $\simeq$ denotes the leading term for large R, the medium of bound pairs is taken into account by the dielectric constant $\tilde{\varepsilon}$, the entropy S is given by $\ln(\Omega/\zeta)$, and Ω is assumed to have the linear extension R. Note that the leading term of the single-particle self-energy is proportional to $\ln(R/r_0)$ in this case [which follows from the

definition given in Sec. II.A.1; compare Eq. (2.5a)]. In the limit $R\to\infty$ the free energy to introduce a single charge is, according to Eq. (2.23), $-\infty$ for $4\tilde{\varepsilon}T>1$ and $+\infty$ for $4\tilde{\varepsilon}T<1$. This may be interpreted as an indication that the neutral Coulomb gas would have no free charges below a critical temperature T_c given by $4\tilde{\varepsilon}(T_c)T_c=1$, whereas free charges would be present above this temperature (Kosterlitz and Thouless, 1973).

The Kosterlitz-Thouless charge-unbinding transition may, of course, be established by more convincing arguments (as will be described further in Secs. II.B.1.a, II.B.1.b, II.D.1, and III). Nevertheless, it is interesting to note that the result $4T_c\tilde{\varepsilon}(T_c)=1$ obtained from Eq. (2.23) remains true for the theoretical treatments so far, at least for fugacities z smaller than some upper limit $z_{\rm up}$ (compare Sec. III.D). One may also note that in the limit $z\to 0$ one has $\tilde{\varepsilon}(T_c)=1$ and $T_c=\frac{1}{4}$, which reflects the trivial fact that the density of Coulomb gas charges and hence polarization vanishes as z goes to zero.

Figure 1 is a sketch of the Kosterlitz-Thouless transition in the (T,z) plane. The Kosterlitz-Thouless critical line starts at $(T,z)=(\frac{1}{4},0)$. The critical temperature T_c diminishes with increasing z because $\tilde{\varepsilon}$ increases and $T_c=1/4\tilde{\varepsilon}$. A two-dimensional Coulomb gas with the fugacity z given as a function of temperature T corresponds to a $z(T)$ trajectory in the (T,z) plane (dashed curve in Fig. 1). The critical temperature T_c for such a Coulomb gas is given by the crossing point between the $z(T)$ trajectory and the Kosterlitz-Thouless critical line (see Fig. 1).

The low-temperature phase does not represent any true long-range order. This lack of long-range order is linked

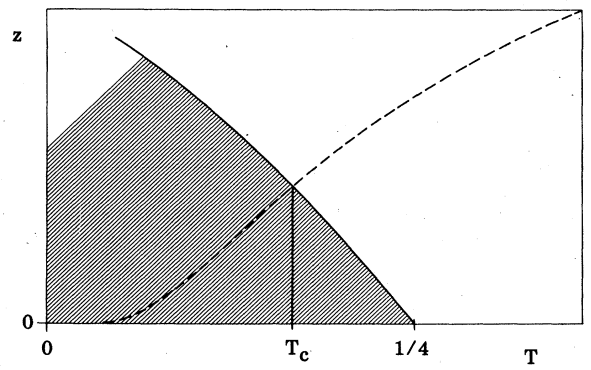

FIG. 1. Sketch of the phase diagram in the (T,z) plane for the neutral two-dimensional Coulomb gas with $\lambda_c=\infty$. The shaded area is the low-temperature "dipole pair" phase. The unshaded area is the high-temperature "free particles + dipole pairs" phase. The two phases are separated by the Kosterlitz-Thouless phase-transition line $4\tilde{\varepsilon}T=1$ (solid curve). The dashed curve is a $z(T)$ trajectory characterizing some specific two-dimensional Coulomb gas. The crossing point between the $z(T)$ trajectory and the Kosterlitz-Thouless phase-transition line defines the Kosterlitz-Thouless critical temperature T_c for the Coulomb gas characterized by this $z(T)$ trajectory.

to the lack of long-range order for a broken continuous symmetry in two dimensions (Mermin and Wagner, 1966; compare the XY model described in Sec. IV.E). However, the charge-density correlations for the low-temperature phase exhibit a "quasi" long-range order in that the charge-density correlation function $\langle\Delta n(r)\Delta n(0)\rangle$ falls off like a power of $1/r$ (Berezinskii, 1971; Kosterlitz and Thouless, 1973), i.e.,

$$\langle\Delta n(r)\Delta n(0)\rangle\sim\begin{cases}\left[\dfrac{1}{r}\right]^{1/T\tilde{\varepsilon}}, & T<T_c\ ,\\ e^{-r/\lambda}, & T>T_c\end{cases}\tag{2.24}$$

(for a derivation see Sec. III.A).

a. Charge unbinding and free-particle density

The basic feature of the screening length λ can be inferred through a heuristic argument which relates the screening length λ to the free-particle density n_F in a self-consistent way (Minnhagen, 1981a; Young and Bohr, 1981).

The screening length λ may be approximately related to n_F by (Minnhagen, 1981a)

$$\lambda^{-2}-\lambda_c^{-2}/\tilde{\varepsilon}=2\pi n_F/(T\tilde{\varepsilon})\ ,\tag{2.25}$$

where λ_c, which is the large-distance cutoff of the particle interaction introduced in Eq. (2.3), has been included for generality [Eq. (2.25) incorporates the feature that $V_L(k=0)=\lambda_c^{-2}$ for $n_F=0$]. The interaction between two Coulomb gas particles, when screened out by the free Coulomb gas charges at large distances, is given by V_L, so that

$$\hat{U}_{\rm eff}(k)=\frac{1}{\tilde{\varepsilon}}\frac{2\pi(\hat{f}_{r_0})^2}{k^2+\lambda^{-2}}\tag{2.26}$$

for small k. The energy needed to create two free charges, of opposite sign and far apart, may then be estimated to be $2|\mu_{\rm eff}|$ where

$$|\mu_{\rm eff}|=-T\ln(z)+\tfrac{1}{2}U_{\rm eff}(0)\ .\tag{2.27}$$

The density of free charges should, accordingly, be given by ($\mu_{\rm eff}=-|\mu_{\rm eff}|$)

$$n_F\simeq\frac{2}{\zeta}e^{\mu_{\rm eff}/T}\ ,\tag{2.28}$$

where the factor 2 reflects the fact that a free particle can have either positive or negative charge. From Eq. (2.26) it follows that $U_{\rm eff}(0)\simeq(1/\tilde{\varepsilon})\ln(\lambda/r_0)$. Consequently, Eqs. (2.26) and (2.28) lead to a self-consistent equation for λ given by

$$\lambda^{-2}-\frac{\lambda_c^{-2}}{\tilde{\varepsilon}}=g\frac{4\pi z}{\tilde{\varepsilon}T}r_0^{-2}(r_0/\lambda)^{1/2\tilde{\varepsilon}T}\ ,\tag{2.29}$$

where the factor $g(z,T)>0$ accounts for the nonleading contributions from $U_{\rm eff}(0)$ (and the difference between ζ and r_0^2). In the limit $\lambda_c\to\infty$, Eq. (2.29) has the solution

(provided $4\pi zg/\tilde{\varepsilon}T<1$ and $r_0/\lambda<1$)

$$\left[\frac{r_0}{\lambda}\right]^2=\left[g\frac{4\pi z}{\tilde{\varepsilon}T}\right]^{1/[1-1/(4\tilde{\varepsilon}T)]} \quad \text{for } T>\frac{1}{4\tilde{\varepsilon}}, \tag{2.30a}$$

$$\lambda^{-2}=0 \quad \text{for } T<\frac{1}{4\tilde{\varepsilon}}. \tag{2.30b}$$

Thus, according to Eqs. (2.30), the Kosterlitz-Thouless transition occurs at a temperature T_c given by $T_c=1/4\tilde{\varepsilon}(T_c^+)$ $[\tilde{\varepsilon}(T_c^+)=\lim_{T\to T_c^+}\tilde{\varepsilon}(T)]$. Furthermore, Eq. (2.30a) gives an expression for λ in terms of z and T. This expression has been verified in the limit of small z for fixed $T>\frac{1}{4}$ (Minnhagen *et al.*, 1978; see Sec. II.D; Young and Bohr, 1981; see Sec. III.C). Equation (2.30a) is also in accord with the renormalization equations for the Kosterlitz-Thouless transition (see Secs. II.D.1 and III.C).

Figure 2 illustrates how the density of free charges n_F vanishes as the Kosterlitz-Thouless transition is approached [as given by Eqs. (2.25) and (2.29) with $4\pi zg/T\tilde{\varepsilon}=$const]. The solid curve is for a finite λ_c and the dashed curve is the $\lambda_c\to\infty$ limit. The divergence between the curves comes when $\lambda_c^{-2}\simeq 2\pi n_F/\tilde{\varepsilon}T$ (arrows in the figure). One notes that the $\lambda_c\to\infty$ limit gives a good approximation of the finite λ_c case for temperatures not too *close* to T_c. Or to put it differently the singular behavior for $\lambda_c=\infty$ at T_c is masked by the finite λ_c. It

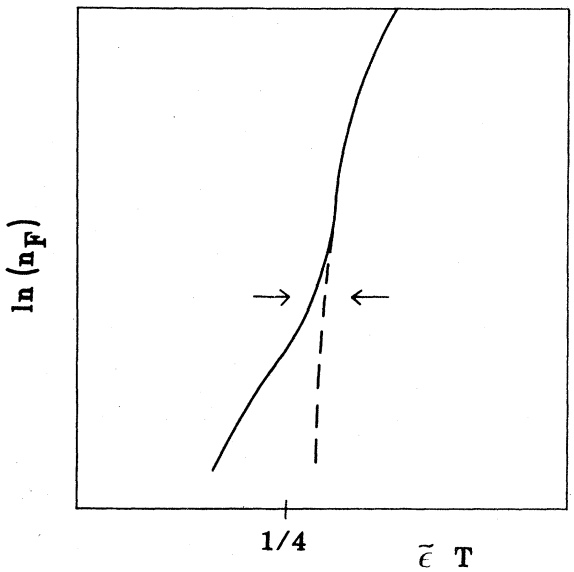

FIG. 2. Sketch of the rapid temperature variation of the density of free charges n_F close to the charge-unbinding transition for a neutral two-dimensional Coulomb gas. The solid curve represents $\ln(n_F)$ as a function of $\tilde{\varepsilon}T$ in the case of a finite λ_c. The dashed curve is the corresponding $\lambda_c=\infty$ case. In the latter case n_F vanishes at $\tilde{\varepsilon}T=\frac{1}{4}$, i.e., at the Kosterlitz-Thouless transition. The arrows indicate where λ_c for the solid curve is approximately equal to the screening length due to free charges, i.e., $\lambda_c^{-2}\simeq 2\pi n_F/\tilde{\varepsilon}T$. [The curves are solutions to Eqs. (2.25) and (2.29).]

should also be noted that, for a large but finite λ_c, there is no phase transition in the sense that the free energy has a nonanalytic behavior, although there is a charge-unbinding transition in the sense that λ has a very rapid variation.

b. *Critical behavior of* $\tilde{\varepsilon}$ *and* λ

It is now reasonable to assume further that the leading temperature dependence for the screening length λ close to T_c comes from the exponent of Eq. (2.30a). In order to extract this leading temperature dependence one needs to know how $\tilde{\varepsilon}(z,T)$ varies close to T_c (for constant z). The following heuristic argument gives a suggestion (Minnhagen, 1981a).

Let us assume that $\tilde{\varepsilon}(T)$ may be Taylor expanded around any $T\neq T_c$ and that the Kosterlitz-Thouless transition is reflected in some nonanalytic behavior at T_c. If we choose the "renormalized" temperature $\bar{T}=\tilde{\varepsilon}T$ as the expansion parameter, then the effect of the bound Coulomb gas pairs is taken into account, so that $\tilde{\varepsilon}$ as a function of $\bar{T}$ might be expected to be better behaved. Expanding around a temperature $\bar{T}_0$ gives

$$\tilde{\varepsilon}(T)=\tilde{\varepsilon}(\bar{T}_0)+\tilde{\varepsilon}'(\bar{T}_0)(\bar{T}-\bar{T}_0)+\frac{\tilde{\varepsilon}''(\bar{T}_0)}{2}(\bar{T}-\bar{T}_0)^2+O((\bar{T}-\bar{T}_0)^3). \tag{2.31a}$$

Directly from the definition of $\bar{T}$ one obtains (by expanding $\tilde{\varepsilon}(\bar{T})=[\bar{T}_0+(\bar{T}-\bar{T}_0)]/[T_0+(T-T_0)]$)

$$\tilde{\varepsilon}(\bar{T})=\tilde{\varepsilon}(\bar{T}_0)+\frac{1}{T_0}(\bar{T}-\bar{T}_0)-\frac{\tilde{\varepsilon}(\bar{T}_0)}{T_0}(T-T_0)+O((\bar{T}-\bar{T}_0)(T-T_0)). \tag{2.31b}$$

These two expansions are compatible [with $\tilde{\varepsilon}''(\bar{T}_0)\neq 0$] only if $\bar{T}-\bar{T}_0\sim T-T_0$ or if $(\bar{T}-\bar{T}_0)^2\sim T-T_0$. The second possibility demands that $\tilde{\varepsilon}'(\bar{T}_0)=1/T_0$, $\tilde{\varepsilon}''(\bar{T}_0^+)<0$, and $\tilde{\varepsilon}''(\bar{T}_0^-)>0$. The first possibility is consistent with $\tilde{\varepsilon}(T)$ having a Taylor expansion around T_0, while the second possibility is not. This may be taken as an indication that the second possibility describes the behavior at the Kosterlitz-Thouless transition. Thus by this hand-waving argument one obtains, close to T_c,

$$\tilde{\varepsilon}(T)=\tilde{\varepsilon}(T_c)\left(1\pm c_{\gtrless}\sqrt{|1-T/T_c|}\right), \quad T-T_c\gtrless 0, \tag{2.32}$$

where $c_{\gtrless}$ are positive constants. The general behavior of $\tilde{\varepsilon}$ for a two-dimensional Coulomb gas with fugacity $z(T)\sim e^{-E_c/T}$ is sketched in Fig. 3. Note that $\tilde{\varepsilon}(T=0)=1$ because $z(T=0)=0$ and that $\tilde{\varepsilon}(T=\infty)=1$ because all pairs break in the high-temperature limit.

Equation (2.32) can be more rigorously motivated by use of renormalization equations for the Kosterlitz-

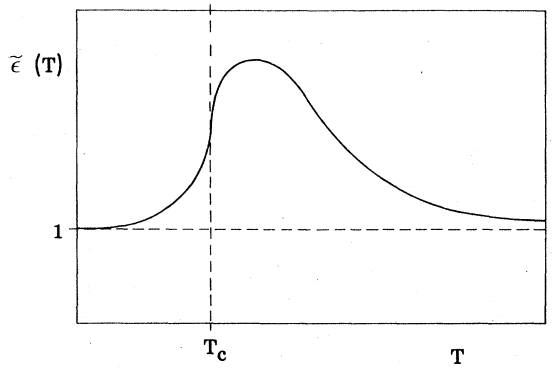

FIG. 3. Sketch of the dielectric constant $\tilde{\varepsilon}(T)$ for a neutral two-dimensional Coulomb gas specified by $z(T)\sim\exp(-E_c/T)$. The figure shows the case when $\tilde{\varepsilon}(T_c^-)=\tilde{\varepsilon}(T_c^+)$. Another possibility would be $\tilde{\varepsilon}(T_c^-)<\tilde{\varepsilon}(T_c^+)$.

Thouless transition (see Sec. III.C). A particularly well-established result is that, below T_c for T_c close enough to $\frac{1}{4}$, one has (Kosterlitz, 1974)

$$\varepsilon(T)=\varepsilon(T_c)(1-c_<\sqrt{|1-T/T_c|}) \text{ with } 4T_c\tilde{\varepsilon}(T_c)=1 .$$

The leading temperature dependence of the exponent in Eq. (2.30a) follows from Eq. (2.32), i.e.,

$$\frac{1}{1-1/(4T\tilde{\varepsilon})}\sim\frac{1}{\sqrt{T-T_c}},\quad T>T_c \tag{2.33}$$

and, accordingly, the leading temperature dependence of the screening length λ, as T_c is approached for $z(T_c)\neq 0$, can be expressed as

$$\lambda\sim\exp\left[\frac{\text{const}}{\sqrt{T-T_c}}\right] . \tag{2.34}$$

This result was first obtained by Kosterlitz (1974) using renormalization equations and will be discussed in more detail in Sec. III.C.

c. Coulomb gas universal jump

The dielectric response function $\varepsilon(r)$ may be defined through the linearly screened potential $V_L(r)$, and the Fourier transforms $\hat{\varepsilon}(k)$ and $\hat{V}_L(k)$ are related by [compare Eqs. (2.21)]

$$\hat{V}_L(k)=\frac{2\pi}{k^2\hat{\varepsilon}(k)}\hat{f}_{r_0}(k) . \tag{2.35}$$

Consequently one finds in the $k\to 0$ limit that [compare Eq. (2.21a)]

$$\frac{1}{\hat{\varepsilon}(0)}=\lim_{k\to 0}\frac{1}{\tilde{\varepsilon}}\frac{1}{1+(\lambda k)^{-2}}=\begin{cases}\dfrac{1}{\tilde{\varepsilon}} & \text{for } T<T_c ,\\ 0 & \text{for } T>T_c\end{cases} . \tag{2.36}$$

This means that a quantity like $1/[T\hat{\varepsilon}(0)]$ jumps from a finite value to zero at T_c. This corresponds to a jump of the superfluid density for a two-dimensional superfluid from a finite value to zero at the critical temperature (Nelson and Kosterlitz, 1977; Minnhagen and Warren, 1981; see Sec. IV.C). For T_c close enough to $\frac{1}{4}$, renormalization equations for the Kosterlitz-Thouless transition lead to $4T_c\tilde{\varepsilon}(T_c)=1$. Consequently the size of the jump for the quantity $1/[T\hat{\varepsilon}(0)]$ at T_c is precisely 4 for this case. This corresponds to the celebrated *universal jump* prediction for a two-dimensional superfluid when described in the Coulomb gas language (Nelson and Kosterlitz, 1977; Minnhagen and Warren, 1981; see Sec. IV.C). However, for a T_c further away from $\frac{1}{4}$, there is some indication that the size of the jump may have a larger, nonuniversal value (Minnhagen, 1985a, 1985b; see Sec. III.D).

d. Free-energy density

The leading z dependence of the free-energy density may also be inferred by the same type of heuristic argument as in Sec. II.B.1.a: the density of free charges n_F corresponds to an effective chemical potential μ_{eff} [see Eqs. (2.27) and (2.28)]. This in turn should correspond to a part of the free-energy density F_F associated with the free charges. This part of the free-energy density is determined by

$$n_F=-\frac{\partial}{\partial\mu_{\text{eff}}}F_F , \tag{2.37}$$

which by Eqs. (2.28) and (2.30a) becomes

$$F_F\simeq-\frac{\tilde{\varepsilon}T^2}{2\pi r_0^2}\left[\frac{4\pi zg}{T\tilde{\varepsilon}}\right]^{2\nu} , \tag{2.38a}$$

where

$$\nu=\frac{2\tilde{\varepsilon}T}{4\tilde{\varepsilon}T-1} . \tag{2.38b}$$

A direct expansion of the partition function Z [Eq. (2.10)] gives, in the limit $\lambda_c\to\infty$, the lowest-order z dependence proportional to z^2 which corresponds to the creation of neutral pairs [compare Eqs. (2.17) and (2.18)]. This gives a contribution to the free-energy density F_B, where $F_B=\text{const}\times z^2$, which may be associated with bound pairs. Consequently the leading z dependence of the free-energy density in the limit of small z is given by

$$F\sim\begin{cases}-\dfrac{T^2}{2\pi r_0^2}\left[\dfrac{4\pi zg}{T}\right]^{2\nu_0}, & T>\frac{1}{2} ,\\ -\dfrac{T}{2}\dfrac{z^2}{\zeta}, & T<\frac{1}{2} ,\end{cases} \tag{2.39a}$$

where

$$\nu_0=\lim_{z\to 0}\nu(z)=\frac{2T}{4T-1} . \tag{2.39b}$$

TABLE I. Some basic features of the two-dimensional Coulomb gas for $\lambda_c \to \infty$.

Screening length	
$\lambda = r_0 \left[\frac{4\pi g z}{\tilde{\varepsilon} T} \right]^{1/[1/(2\tilde{\varepsilon}T)-2]}, \quad T > T_c$	See Eq. (2.30a)
$\lambda \sim r_0 e^{\text{const}/\sqrt{T-T_c}}, \quad T \to T_c^+$	See Eq. (2.34)
$\lambda = \infty, \quad T < T_c^+$	See Eq. (2.30)
Dielectric constant due to bound pairs	
$\tilde{\varepsilon} = \frac{1}{4T_c}(1 + c_> \sqrt{\lvert 1 - T/T_c \rvert}), \quad T \to T_c^+$	See Eq. (2.32)
$\tilde{\varepsilon} = \frac{1}{4T_c}(1 - c_< \sqrt{\lvert 1 - T/T_c \rvert}), \quad T \to T_c^-$	See Eq. (2.32)
Universal jump	
$1/[T\hat{\tilde{\varepsilon}}(0)] = 0, \quad T \to T_c^+$	See Eq. (2.36)
$1/[T\hat{\tilde{\varepsilon}}(0)] = 4, \quad T \to T_c^-$	See Eq. (2.36)
Free-energy density	
$F \sim -\frac{T^2}{4\pi r_0^2} \left[\frac{4\pi z g}{T} \right]^{2T/(4T-1)}, \quad T > \frac{1}{2}, \quad z \to 0$	See Eq. (2.39)
$F \sim -\frac{Tz^2}{2\zeta}, \quad T < \frac{1}{2}, \quad z \to 0$	See Eq. (2.39)

[Equation (2.39b) follows because $\lim_{z\to 0}\tilde{\varepsilon}(z) = 1$.] This result for F was first deduced by Zittartz (1976) and was shown more rigorously by Minnhagen *et al.* (1978; see Sec. II.D).

The physical interpretation of Eqs. (2.39) is that in the limit $z \to 0$ the free charges dominate for $T > \frac{1}{2}$. This is also reflected in the equation of state: the pressure is given by $p = T(\partial/\delta\Omega)\ln(z) = -F$. From Eq. (2.37) [with (μ, n) instead of (μ_{eff}, n_F)] and Eq. (2.39) it then follows that

$$p = n(T - \tfrac{1}{4}) \quad \text{for } T > \tfrac{1}{2}, \tag{2.40a}$$

$$p = \tfrac{1}{2}nT \quad \text{for } T < \tfrac{1}{2}, \tag{2.40b}$$

where n is the total density of charges. One notes that Eq. (2.40b) is the equation of state for an ideal gas of neutral pairs with density $n/2$, whereas the nT term on the right-hand side of Eq. (2.40a) indicates a gas with the free-particle density n. The equation of state given by Eqs. (2.40) was derived by Hauge and Hemmer (1971) in the limiting case $r_0 \to 0$. The limit $r_0 \to 0$ with z fixed is equivalent to the limit $z \to 0$ with r_0 fixed (Minnhagen *et al.*, 1978). In other words, results obtained for the point-charge model reflect the small-z limit of the models in which a single charge has a finite extension.

In this section the general features of the two-dimensional neutral Coulomb gas have been described and motivated by heuristic reasoning. Some of the most basic results are summarized in Table I. We shall come back to these results in Secs. II.D.1, III.C, and III.D.

2. Non-neutral Coulomb gas

In this section we generalize the reasoning for the screening length λ given in Sec. II.B.1.a to the case of a non-neutral two-dimensional Coulomb gas and describe the basic features.

For the non-neutral Coulomb gas model the creation energies of positive and negative charge and unequal and the average charge density $\langle \Delta n \rangle = (1/\Omega)\langle N_+ - N_- \rangle$ is nonzero. The energy needed to introduce an additional positive charge into the non-neutral Coulomb gas may be estimated in the following way (Minnhagen, 1981a).

First introduce the additional charge uniformly distributed over the two-dimensional volume Ω. This requires an overall charging energy given by

$$\int d^2r\, U(\mathbf{r})\langle \Delta n \rangle = \hat{U}(k=0)\langle \Delta n \rangle = 2\pi\lambda_c^2 \langle \Delta n \rangle . \tag{2.41}$$

The next step is to assemble this uniform charge into a single charge distribution $f_{r_0}(r)$. In analogy with the neutral Coulomb gas, this involves the self-energy $\frac{1}{2}U_{\text{eff}}(0)$ [compare Eqs. (2.26) and (2.27)]. The nonelectrostatic energy needed to create a positive (negative) Coulomb gas particle is $E_{c\,(-)}^{+}\Delta E = -T\ln(z)_{(-)}^{+}\Delta E$ [compare Eq. (2.7)]. It follows that the total energy required to create a positive (negative) charge is

$$\mu_{\text{eff}}^{(+)} = -T\ln(z)_{(-)}^{+}\Delta E + \tfrac{1}{2}U_{\text{eff}}(0)_{(-)}^{+}2\lambda_c^2\langle \Delta n \rangle , \tag{2.42}$$

and the corresponding density of free positive (negative) charges is

$$n_F^{(+)} \simeq \frac{e^{\mu_{\rm eff}^{(+)}/T}}{\zeta} . \tag{2.43}$$

Equations (2.42) and (2.43) lead (by using the expression $n_F^2 - \langle \Delta n \rangle^2 = 4 n_F^+ n_F^-$ to eliminate ΔE and $\lambda_c^2 \langle \Delta n \rangle$), in analogy with Eq. (2.29) for the neutral Coulomb gas, to a self-consistent equation for n_F which, in the limit $\lambda/\lambda_c \to 0$, may be expressed as (Minnhagen, 1981a)

$$|\langle \Delta n \rangle| = n_F \left[1 - (z\tilde{g})^2 \left[\frac{2\pi n_F r_0^2}{\tilde{\varepsilon} T} \right]^{1/2\tilde{\varepsilon}T - 2} \right]^{1/2} , \tag{2.44}$$

where $\langle \Delta n \rangle = n_F^+ - n_F^-$, $n_F = n_F^+ + n_F^-$, $\tilde{g} = (4\pi/\tilde{\varepsilon}T) g$, and $g(z, T, \Delta E)$ takes the nonleading terms of $U_{\rm eff}(0)$ into account [compare Eq. (2.29)]. Equation (2.44) describes a crossover from a low-temperature regime $n_F \simeq \langle \Delta n \rangle$ to a high-temperature regime where the dominant contribution to n_F is generated by charge unbinding. This is illustrated in Fig. 4, which is a solution of Eq. (2.44) with constant $\langle \Delta n \rangle$ and $z\tilde{g}$ (solid curve in the figure).

Equations (2.25), (2.42), and (2.43) may also be combined into (Minnhagen, 1981a)

$$\langle \Delta n \rangle = n_F \tanh \left[\frac{\Delta E}{T} - \frac{2\pi \lambda_c^2}{T} \langle \Delta n \rangle \right] , \tag{2.45}$$

which means that in the limit $\lambda_c \to \infty$

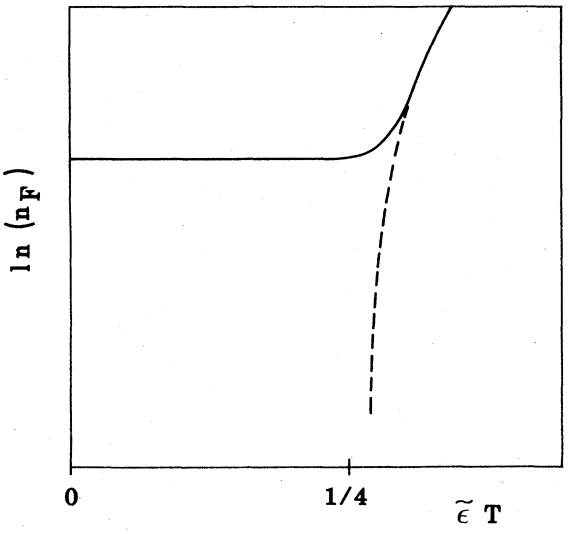

FIG. 4. The density of free charges n_F for a non-neutral two-dimensional Coulomb gas close to the charge-unbinding transition. Solid curve represents $\ln(n_F)$ as a function of $\tilde{\varepsilon}T$ obtained from Eq. (2.44) for $\langle \Delta n \rangle \neq 0$ and $z\tilde{g}$ constant. The basic feature is a crossover from a low-temperature region $T < \frac{1}{4}$ with $n_F \simeq \langle \Delta n \rangle$ to a high-temperature region $T > \frac{1}{4}$ with $n_F \gg \langle \Delta n \rangle$. The dashed curve represents the $\langle \Delta n \rangle = 0$ case, i.e., the corresponding result for the neutral two-dimensional Coulomb gas.

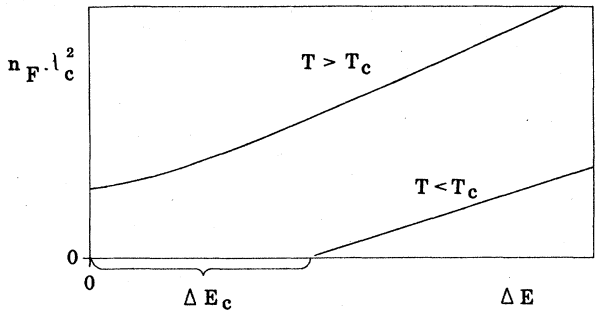

FIG. 5. Density of free Coulomb gas charges n_F as a function of the difference in energy needed to create a positive and a negative Coulomb gas charge. The two solid curves represent $n_F \lambda_c^2$ as a function of ΔE ($2\Delta E$ = the difference in energy needed to create a positive and a negative Coulomb gas charge), illustrating the two cases $T > T_c$ and $T < T_c$, respectively. The $T < T_c$ case has a "quasi" critical behavior in that the density of free Coulomb gas charges is almost zero up to a "quasi" critical ΔE_c.

$$\Delta E \simeq 2\pi \lambda_c^2 \langle \Delta n \rangle \tag{2.46}$$

is a necessary condition for a nonzero $\langle \Delta n \rangle$. In Fig. 5 the density of free charges is plotted as a function of ΔE. Below T_c the density of free charges [as determined by Eqs. (2.25), (2.42), and (2.43)] has a quasicritical behavior in the sense that n_F remains almost zero up to a characteristic ΔE_c where ΔE_c is approximately given by (Minnhagen, 1981a)

$$\Delta E_c \simeq \frac{1 - 4T\tilde{\varepsilon}}{2\tilde{\varepsilon}} \ln \left[\frac{\lambda_c}{r_0} \right] . \tag{2.47}$$

This ΔE_c is related to a "quasicritical" H_{c1} for superconducting films (Doniach and Huberman, 1979; Minnhagen, 1981a).

For low enough temperatures the positive (negative) charges will dominate the free energy for $\langle \Delta n \rangle \gtrless 0$. Such a one-component Coulomb gas forms a Wigner lattice for small enough temperatures (Wigner, 1938; Abrikosov, 1957). This will be discussed somewhat further in connection with superconducting films (see Sec. VII.B.2).

3. Neutral Coulomb gas in an electric field

The two-dimensional neutral Coulomb gas has no free charges below T_c for $\lambda_c = \infty$ (see Sec. II.B.1.a). However, pair breaking may be induced below T_c by applying a constant external electric field. This is analogous to applying a constant current to a superconducting film or to a finite flow velocity in the case of superfluid films [e.g., compare Eq. (4.62) in Sec. IV.B.2]. The density of free charges, n_F generated by the external electric field $\mathbf{D}$ can be estimated by the same type of heuristic reasoning as in Secs. II.B.1.a and II.B.2.

The energy needed to create a positive and negative charge at positions $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively, is given by [compare Eq. (2.27)]

$$2|\mu_{\rm eff}(\mathbf{r}_1,\mathbf{r}_2)|=2E_c+U_{\rm eff}(0)-U_{\rm eff}(r_{12})+\frac{1}{\tilde{\varepsilon}}(\mathbf{D}\cdot\mathbf{r}_2-\mathbf{D}\cdot\mathbf{r}_1). \tag{2.48}$$

If the force acting between the charges, F_{12}, is repulsive, then the pair is broken into two free charges. The minimum energy needed to create a broken pair is hence determined by the condition $F_{12}=0$. Obviously the minimum energy corresponds to a dipole pair oriented parallel to the field and the condition

$$-\frac{d}{dr_{12}}U_{\rm eff}(r_{12})=\frac{1}{\tilde{\varepsilon}}D \tag{2.49a}$$

or

$$\frac{1}{r_{12}}\simeq D \tag{2.49b}$$

where in the last step the approximation $U_{\rm eff}(r)\simeq-(1/\tilde{\varepsilon})\ln(r/\lambda)$ was used [compare Eqs. (2.5) and (2.26)]. The effective chemical potential for creating a free charge is accordingly

$$\mu_{\rm eff}\simeq -E_c+\tfrac{1}{2}[U_{\rm eff}(D^{-1})-U_{\rm eff}(0)]+\frac{1}{\tilde{\varepsilon}}, \tag{2.50}$$

and, the density of free charges

$$n_F\simeq\frac{1}{\zeta}e^{\mu_{\rm eff}/T}.$$

Using the same estimate as for Eq. (2.49b), i.e., $U_{\rm eff}(r)\simeq-(1/\tilde{\varepsilon})\ln(r/\lambda)$, gives

$$n_F\sim(Dr_0)^{1/2T\tilde{\varepsilon}}. \tag{2.51}$$

This means that the density of free charges generated by an external electric field goes like a power of the field. The result was first obtained by Ambegaokar *et al.* (1978) and Myerson (1978) for the analogous case of superfluid films.

C. Sine-Gordon formulation

The results for the two-dimensional Coulomb gas given in Sec. II.B (see Table I) were motivated by intuitive reasoning. An obvious question is to what extent these results can be verified by more stringent methods. Some aspects of this question will be treated in this section and the ensuing sections II.D and III.

Here we describe how the model of the two-dimensional Coulomb gas may be transformed into a sine-Gordon field theory. The advantage of this transformation is that it makes available to us standard field theory diagrammatics and renormalization procedures. This opens up a systematic way of obtaining results for the two-dimensional Coulomb gas.

First let us examine the precise connection between the two-dimensional Coulomb gas and the sine-Gordon field theory. This connection rests on the fact that the partition function Z for the Coulomb gas can be expressed as a functional integral over a real field (Fröhlich, 1976; Polyakov, 1977; Samuel, 1978). As described in Sec. II.B.1, the charge-unbinding transition may be characterized in terms of the screening length λ and the dielectric constant $\tilde{\varepsilon}$. The field theory definitions of these quantities will be given. In Sec. II.D.1 we then describe how these field theory definitions may be used to generate systematic expansions.

The key quantities in the field theory formulation are correlation functions of the form

$$\langle O[\varphi(\mathbf{r})]\rangle_m = \frac{\int d\varphi\exp\left[-\int d^2r\left[\frac{(\nabla\varphi)^2}{2}+\frac{m^2\varphi^2}{2}\right]\right]O[\varphi(r)]}{\int d\varphi\exp\left[-\int d^2r\left[\frac{(\nabla\varphi)^2}{2}+\frac{m^2\varphi^2}{2}\right]\right]} \tag{2.52}$$

where $\varphi(\mathbf{r})$ is a real field, $O[\varphi]$ is a functional of this field, and the integrals are functional integrals with respect to the field $\varphi(\mathbf{r})$. The key to the transformation between the Coulomb gas and the sine-Gordon theory is the correlation function

$$\left\langle\exp\left[-i\sqrt{2\pi/T}\sum_i s_i\tilde{\varphi}(\mathbf{r}_i)\right]\right\rangle_{m=\lambda_c^{-1}} =\exp\left[-\frac{1}{2T}\sum_{ij}s_is_jU(r_{ij})\right], \tag{2.53}$$

where the right-hand side is obtained by straightforwardly performing the functional integrals implied by $\langle\ \rangle_{m=\lambda_c^{-1}}$ on the left-hand side. $U(r)$ is given by Eq. (2.4) and $\tilde{\varphi}(\mathbf{r})$ by

$$\tilde{\varphi}(\mathbf{r})=\int d^2r' f_{r_0}(|\mathbf{r}-\mathbf{r}'|)\varphi(\mathbf{r}'). \tag{2.54}$$

Inserting the identity given by Eq. (2.53) into the definition of the partition function Z [see Eq. (2.10)] gives

$$Z=\left\langle\exp\left[\frac{\int d^2r}{\zeta}\{\bar{z}_+\exp[i\sqrt{2\pi/T}\,\tilde{\varphi}(\mathbf{r})]+\bar{z}_-\exp[-i\sqrt{2\pi/T}\,\tilde{\varphi}(\mathbf{r})]\}\right]\right\rangle_{m=\lambda_c^{-1}}, \tag{2.55}$$

where $\bar{z}_{\substack{+\\(-)}}$ differ from the "true" one-particle fugacity [see Eq. (2.10)] by the one-particle electrostatic self-energy, i.e.,

$$\bar{z}_{\substack{+\\(-)}} = z_{\substack{+\\(-)}} \exp\left[\frac{1}{2T}U(0)\right]$$

$$= \exp[-(E_{c(\substack{+\\-})}\Delta E)/T] \ . \tag{2.56}$$

For the neutral Coulomb gas in the limit $\lambda_c = \infty$ the field theory expression for Z reduces to

$$Z = \left\langle \exp\left[\frac{\int d^2r}{\zeta} 2z \cos[\sqrt{2\pi/T}\,\tilde{\varphi}(\mathbf{r})]\right]\right\rangle_{m=\lambda_c^{-1}=0} \tag{2.57}$$

with $z = e^{-E_c/T}$. This means that Z for the neutral Coulomb gas with $\lambda_c = \infty$ may alternatively be expressed as

$$Z = \frac{\int d\varphi \exp\left[-\int d^2r\, H_{\rm sG}\right]}{\int d\varphi \exp\left[-\int d^2r (\nabla\varphi)^2/2\right]} \ , \tag{2.58a}$$

where

$$H_{\rm sG} = \frac{[\nabla\varphi(r)]^2}{2} - \frac{2z}{\zeta}\cos[\sqrt{2\pi/T}\,\tilde{\varphi}(\mathbf{r})] \tag{2.58b}$$

is the Hamiltonian density for a sine-Gordon (sG) field theory. A correlation function within the sine-Gordon field theory is given by

$$\langle O[\varphi]\rangle^{\rm sG} = \frac{\int d\varphi \exp\left[-\int d^2r\, H_{\rm sG}\right] O[\varphi]}{\int d\varphi \exp\left[-\int d^2r H_{\rm sG}\right]} \ , \tag{2.59}$$

where $O[\varphi]$ is some functional of the field φ and the superscript sG is introduced in order to distinguish sine-Gordon correlation functions from the thermal average of the Coulomb gas $\langle\ \rangle$ and the correlation functions $\langle\ \rangle_m$ defined by Eq. (2.52).

A sine-Gordan correlation function of special interest in the present context is the one-particle Green's function $G(r)$ defined by

$$G(r) = \langle \varphi(r)\tilde{\varphi}(0)\rangle^{\rm sG} \ . \tag{2.60}$$

It is related to the linearly screened potential V_L by (see, for example, Minnhagen, 1985b)

$$V_L(r) = 2\pi G(r) \ . \tag{2.61}$$

The Green's function G is related to the one-particle irreducible self-energy Σ by

$$\hat{G}(k) = \frac{\hat{f}_{r_0}(k)}{k^2 - \hat{\Sigma}(k)} \ . \tag{2.62}$$

From Eqs. (2.61) and (2.62), together with the identifications

$$\tilde{\varepsilon} = 1 - \frac{\partial}{\partial k^2}\Sigma(k)\bigg|_{k=0} \ , \tag{2.63a}$$

$$\lambda^{-2} = -\frac{\Sigma(0)}{1 - \frac{\partial}{\partial k^2}\Sigma(k)}\bigg|_{k=0} \ , \tag{2.63b}$$

one regains Eq. (2.21a),

$$\hat{V}_L(k) = \frac{2\pi\hat{f}_{r_0}(k)}{\tilde{\varepsilon}(k^2+\lambda^{-2})} \quad \text{for } k \ll 0 \ , \tag{2.64}$$

which is the definition of $\tilde{\varepsilon}$ and λ (see Sec. II.B.1). This means that the key quantities $\tilde{\varepsilon}$ and λ for the characterization of the charge-unbinding transition are, within the field theory formulation, defined by Eqs. (2.63a) and (2.63b), respectively. As a consequence they may be calculated by standard field theory diagrammatics and renormalization procedures (as will be discussed somewhat further in Sec. II.D.1).

Another example of the mapping between the Coulomb gas and the sine-Gordon formulation is given by the charge-density correlation function: the point-charge density for a configuration of N particles is defined by

$$\Delta n^0(r) = \sum_i^N s_i \delta(\mathbf{r} - \mathbf{r}_i) \ , \tag{2.65}$$

where $\mathbf{r}_i$ denotes the position of the charge i. The point-charge-density correlation function may be expressed as (Minnhagen, 1985b)

$$\langle \Delta n^0(r)\Delta n^0(0)\rangle$$

$$= -\frac{4z^2}{\zeta^2}\langle \sin[\sqrt{2\pi/T}\,\tilde{\varphi}(\mathbf{r})]\sin[\sqrt{2\pi/T}\,\tilde{\varphi}(0)]\rangle^{\rm sG}$$

$$+ \delta(\mathbf{r})\frac{2z}{\zeta}\langle \cos[\sqrt{2\pi/T}\,\tilde{\varphi}(0)]\rangle^{\rm sG} \ . \tag{2.66}$$

It may also be noted that the sine-Gordon Hamiltonian density $H_{\rm sG}$ [Eq. (2.58b)] reduces to a Klein-Gordon (KG) Hamiltonian density in the high-T limit, i.e.,

$$H_{\rm sG} = H_{\rm KG} = \frac{(\nabla\varphi)^2}{2} + \frac{4\pi z}{\zeta T}\tilde{\varphi}^2 \quad \text{for } T \gg 1 \ . \tag{2.67}$$

In other words, the high-T limit, which for a charged gas corresponds to a Debye-Hückel limit, is associated with the Klein-Gordon limit of the field theory formulation.

D. Expansion approaches

In this section we sketch how some of the results given in Sec. II.B (see Table I) may be obtained by more systematic expansions. The sine-Gordan formulation has turned out to be especially useful in this respect, and we shall for this reason concentrate the presentation on expansions within this field theory formulation (Sec. II.D.1). It is, of course, also possible to define expansion procedures directly in the usual statistical mechanical formulation of the two-dimensional Coulomb gas, e.g., various

types of cluster expansions. Progress made for the two-dimensional Coulomb gas using these, from a statistical mechanical point of view, more conventional methods is briefly mentioned in Sec. II.D.2.

1. Sine-Gordon expansion

As described in Sec. II.C, the screening length λ and the dielectric constant $\tilde{\epsilon}$ are related to the irreducible self-energy of the Green's function for the sine-Gordon theory [compare Eqs. (2.60), (2.61), and (2.63)]. This means that λ and $\tilde{\epsilon}$ can be calculated through systematic expansions of the irreducible self-energy. This section describes such an expansion. It also contains a brief description of what is involved in constructing the renormalization-group equations for the sine-Gordon theory.

In order to define an expansion one may split the sine-Gordon Hamiltonian density into two parts, $H_{\rm sG}=H_0+H_I$, where

$$H_0(r)=\frac{[\nabla\varphi(r)]^2}{2}+\frac{\lambda^{-2}}{2}[\tilde{\varphi}(r)]^2 , \tag{2.68a}$$

$$H_I(r)=-\frac{2z}{\zeta}\cos[\sqrt{2\pi/T}\,\tilde{\varphi}(\mathbf{r})]-\frac{\lambda^{-2}}{2}[\tilde{\varphi}(r)]^2 . \tag{2.68b}$$

The expansion of the Green's function G is then defined in terms of a "bare" Green's function $G^0(r)$ given by [compare Eq. (2.52)]

$$G^0(r)=\langle\varphi(r)\tilde{\varphi}(0)\rangle_{m=\lambda^{-1}} . \tag{2.69}$$

By expanding the Green's function G in the fugacity z one obtains a sequence of Feynman diagrams (for details see, for example, Samuel, 1978). The first-order diagrams are drawn in Fig. 6. For example, the value of the first diagram in the sequence is

$$\frac{4\pi z}{T}\int\frac{d^2r'}{\zeta}G^0(|\mathbf{r}-\mathbf{r}'|)G^0(r') .$$

The sequence of first-order diagrams is, from a structural point of view, a sequence of tadpole diagrams. The sine-Gordon theory has the property that any sequence of tadpole diagrams when summed gives a factor $e^{-(\pi/T)G^0(0)}$ (Coleman, 1975). Thus the sum of all the first-order diagrams is

$$\frac{4\pi z}{T}e^{-(\pi/T)G^0(0)}\int\frac{d^2r'}{\zeta}G^0(|\mathbf{r}-\mathbf{r}'|)G^0(r') .$$

A general way of stating this property of the sine-Gordon theory is the following: All tadpole diagrams are taken into account by "renormalizing" z to $\tilde{z}=ze^{-(\pi/T)G^0(0)}$ (Coleman, 1975). After such a vertex renormalization only one first-order diagram remains, as illustrated in Fig. 6.

The irreducible self-energy to first order in z is trivially obtained by removing the two external G^0 legs from the renormalized first-order diagram, i.e.,

FIG. 6. First-order Feynman diagrams for the sine-Gordon one-particle Green's function. This sequence of diagrams is equivalent to one single diagram with a renormalized vertex (see text).

$$\Sigma^{(1)}(\mathbf{r},\mathbf{r}')=\frac{4\pi z}{\zeta T}e^{-(\pi/T)G^0(0)}\delta(\mathbf{r}-\mathbf{r}') \tag{2.70a}$$

or

$$\hat{\Sigma}^{(1)}(k)=\frac{4\pi z}{T\zeta}e^{-(\pi/T)G^0(0)} . \tag{2.70b}$$

By inserting $\Sigma^{(1)}$ into Eqs. (2.63) and observing that the leading dependence of $G^0(0)$ for large λ is $G^0(0)\simeq(1/2\pi)\ln(\lambda/r_0)$, one obtains

$$\left[\frac{r_0}{\lambda}\right]^2=\left[g(T)\frac{4\pi z}{T}\right]^{1/(1-1/4T)} \quad \text{for } T>\tfrac{1}{4} \tag{2.71a}$$

and

$$\tilde{\epsilon}=1 , \tag{2.71b}$$

where $g(T)$ accounts for the nonleading contributions [compare Eq. (2.29)]. This means that Eqs. (2.30) are regained by sine-Gordon diagrammatics in the limit of small z. By considering higher-order diagrams for the irreducible self-energy it becomes plausible that Eqs. (2.71) remain correct to all orders; the higher-order terms will only produce a new $g(T)$ (Minnhagen *et al.*, 1978).

Using this type of sine-Gordon diagrammatics one can also show that the free-energy density in the small-z limit is given by (Minnhagen *et al.*, 1978)

$$F=F_F+F_B , \tag{2.72a}$$

$$F_F=-\frac{T^2}{4\pi\nu_0\zeta}\left[\frac{4\pi z}{T}\right]^{2\nu_0}\left[c_0+O\left[\left[\frac{4\pi z}{T}\right]^{2\nu_0}\right]\right] , \tag{2.72b}$$

$$F_B=-\frac{T^2}{2\pi\zeta}\sum_{p=1}^{\infty}e_{2p}\left[\frac{4\pi z}{T}\right]^{2p} , \tag{2.72c}$$

where $2\nu_0=(1-1/4T)^{-1}$, $c_0=1+O(T^{-3})$, and e_{2p} are functions of T, e.g., $e_2=T^2/2(1-2T)$. At the "special" temperature $T_p=p/2(2p-1)$ the pth term is proportional to $z^{2p}\ln(z)$. These results for the free energy where first hypothesized by Zittartz (1976) on the basis of scaling arguments. It should also be noted that in the small-z limit the results obtained from sine-Gordon diagrammatics given by Eq. (2.72) confirm and extend the heuristically obtained results given by Eq. (2.39).

The sine-Gordon diagrammatics sketched above are defined in terms of a bare Green's function G^0 [see Eq.

(2.69)] and a renormalized mass λ^{-1} [see Eq. (2.68)]. However, this does not turn out to be a useful expansion close to and below the critical line. The $g(T)$ in Eq. (2.71a) diverges as $g(T)\sim 1/(T-\frac{1}{4})$ when $T=\frac{1}{4}$ is approached (Minnhagen *et al.*, 1978). This means that the z interval for which the diagrammatic expansion is valid shrinks to zero at $T=\frac{1}{4}$. Nevertheless it is possible to construct useful expansions close to the critical line by use of renormalization-group theory (see, for example, Amit, 1984). For the sine-Gordon theory it turns out that the logarithmic infinities arising in the diagrammatic expansion at the critical line for $r_0\to 0$ can be absorbed in two renormalization constants—a wave-function renormalization and a coupling-constant renormalization. The corresponding renormalization-group equations can then be constructed. We shall here only briefly sketch what is involved in such a procedure:

The renormalization procedure may be defined in terms of two renormalization constants ε_R and c_R, together with the corresponding renormalized quantities

$$\varphi_R=\sqrt{\varepsilon_R}\phi,\quad \frac{z_R}{T}=c_R\frac{z}{T},$$
$$T_R=\sqrt{\varepsilon_R}T,\quad \lambda_R=\sqrt{\varepsilon_R}\lambda_c .$$

Obviously the Green's function G can be expressed as

$$G=\frac{1}{\varepsilon_R}\langle\varphi_R(r)\varphi_R(0)\rangle^{\mathrm{sG}},$$

so that

$$\hat{G}=\frac{1}{\varepsilon_R}\frac{1}{k^2+\dfrac{\lambda_c^{-2}}{\varepsilon_R}-\hat{\Sigma}_R(k)} \tag{2.73}$$

and consequently

$$\lambda_R^2\hat{\Sigma}_R(k=0;z_R/T,T_R,\lambda_R)=\lambda_c^2\hat{\Sigma}(k=0;z/T,T,\lambda_c), \tag{2.74a}$$

$$\frac{d}{dk^2}\hat{\Sigma}_R(k;z_R/T,T_R,\lambda_R)\bigg|_{k=0}=\frac{d}{dk^2}\frac{1}{\varepsilon_R}\hat{\Sigma}(k;z/T,T,\lambda_c)\bigg|_{k=0}+(\varepsilon_R^{-1}-1), \tag{2.74b}$$

$$\frac{d}{dk^2}\hat{\Sigma}(k;z_R/T,T_R,\lambda)\bigg|_{k=0}=0 . \tag{2.74c}$$

The last equation is enforced so as to ensure that the physical value of ε_R is given by $\varepsilon_R=\tilde{\varepsilon}$ [compare Eqs. (2.63) and (2.64)].

The renormalization constants can be defined so as to make the right-hand sides of Eqs. (2.74a) and (2.74b) finite, order by order, in a double expansion in the parameters z_R/T and $T_R-\frac{1}{4}$ (Amit *et al.*, 1980). As an example let us consider Σ to lowest order [Eq. (2.70b)],

$$\lambda_R^2\hat{\Sigma}_R(k=0)=\frac{\lambda_c^2 4\pi z_R}{Tc_R\zeta}\exp\left[-\frac{\pi\varepsilon_R}{T_R}G^0(0)\right]\sim\frac{1}{c_R}\left[\frac{\lambda_c}{r_0}\right]^{2-\varepsilon_R/2T_R}, \tag{2.75a}$$

$$\frac{d}{dk^2}\Sigma_R(k;z_R/T,T_R,\lambda_R)\bigg|_{k=0}=\varepsilon_R^{-1}-1=0 . \tag{2.75b}$$

The first of these equations becomes logarithmically divergent for $r_0\to 0$ at the point where the exponent changes from negative to positive. This logarithmic divergency may be absorbed into c_R by

$$c_R=1+\left[2-\frac{\varepsilon_R}{2T_R}\right]\ln(\lambda_c/r_0) . \tag{2.76}$$

The other renormalization constant, ε_R, is according to Eq. (2.75b) equal to 1. To lowest order in z_R/T and $T_R-\frac{1}{4}$ this renormalization leads to a differential equation for z_R in the variable $\tilde{l}=\ln(\lambda_c/r_0)$ (see, for example, Amit, 1984), i.e.,

$$\frac{d}{d\tilde{l}}z_R(\tilde{l})=\left[2-\frac{1}{2T_R(\tilde{l})}\right]z_R(\tilde{l}) . \tag{2.77a}$$

The corresponding equation for $T_R(\tilde{l})$ (i.e., the equation to same order in the double expansion in the parameters z_R/T and $T_R-\frac{1}{4}$) can be obtained by considering the next higher-order diagrams for Σ. It is given by

$$\frac{d}{d\tilde{l}}\frac{1}{T_R(\tilde{l})}=-\frac{2\pi^2}{T^2}z_R^2(\tilde{l}) . \tag{2.77b}$$

Equations (2.77) are the Kosterlitz renormalization-group (RG) equations, which were first derived by Kosterlitz (1974) within the Coulomb gas formulation. They have been rederived within the sine-Gordon formulation by Ohta (1978), Wiegman (1978), and Amit *et al.* (1980).

The procedure sketched above has been carried out to one order higher in the double expansion in the variables z_R/T and $T_R-\frac{1}{4}$ leading to (Amit *et al.*, 1980)

$$\frac{d}{d\tilde{l}}[z_R(\tilde{l})]=\frac{z_R(\tilde{l})}{2}\left[4-\frac{1}{T_R(\tilde{l})}\right]+\frac{5}{4}\frac{\pi^2}{T^2}z_R^3(\tilde{l}) , \tag{2.78a}$$

$$\frac{d}{d\tilde{l}}\left[\frac{1}{T_R(\tilde{l})}\right]=-\frac{2\pi^2 z_R^2(\tilde{l})}{T^2}+\frac{\pi^2}{T^2}z_R^2(\tilde{l})\left[4-\frac{1}{T_R(\tilde{l})}\right] . \tag{2.78b}$$

It turns out that the values of the coefficients in front of the third-order terms in Eqs. (2.78) (third order in the double expansion in the variables z_R and $T_R-\frac{1}{4}$) are nonuniversal, whereas the coefficients in front of the second-order terms are universal. This means that the values of the coefficients in front of the third-order terms depend on the explicit choice of the charge distribution function f_{r_0}, whereas the coefficients in front of the second-order terms are independent of this choice.

The physical quantity $\tilde{\varepsilon}=\varepsilon_R(\infty)$ may be obtained by integrating the equations from $\tilde{l}=0$ to $\tilde{l}=\infty$ starting from the initial conditions $z_R(0)=0$ and $T_R(0)=T$ [compare Eqs. (2.75b) and (2.76)]. The guaranteed validity of this calculation for $\tilde{\varepsilon}$ is by construction restricted to the parameter range $z \ll 1$ and $0<\frac{1}{4}-T \ll 1$. The renormalization-group equations will be discussed in more detail in Sec. III.

2. Cluster expansions

A more traditional way of constructing expansions for a statistical mechanical model like the two-dimensional Coulomb gas is through Mayer cluster expansion techniques (see, for example, Friedman, 1962). Some of the heuristic results presented in Sec. II.B have also been obtained through these types of expansions. For example, the equation of state for $z \to 0$ [Eqs. (2.40)] was obtained by Hauge and Hemmer (1971; see also Deutsch and Lavaud, 1974); the screening length λ [Eq. (2.71a) for $T>\frac{1}{2}$ in the limit $z \to 0$ was obtained by Everts and Koch (1977)]; the Kosterlitz renormalization-group equations [Eqs. (2.77)] were rederived using Mayer cluster expansion techniques by Høye and Olaussen (1980, 1981). A review of these expansion techniques for the two-dimensional Coulomb gas is, however, deemed outside the scope of the present article.

E. Comment on fermion analogies

In Sec. II.D we briefly indicated how information on the Coulomb gas can be obtained by expansion approaches within the sine-Gordon field theory formulation and furthermore how this formulation lends itself to renormalization-group procedures. For completeness it was mentioned that some results have also been obtained by expansion procedures directly in the Coulomb gas formulation. For example, the Kosterlitz renormalization-group equations were originally derived within the Coulomb gas formulation (Kosterlitz, 1974).

An important fact for the transformation between the Coulomb gas model and the sine-Gordon formulation described in Sec. II.C is that it is mathematically exact. There exist yet other transformations of the two-dimensional Coulomb gas model which connect it to one-dimensional fermion models. The existence of such connections is in itself of great interest. The transformations are, however, not entirely exact (as will be explained further below), and because of this they are somewhat less useful in the present context and will consequently not be reviewed in this article. At any rate these fermion analogies form a subject by themselves.

One route to establishing a fermion connection is to start from the Coulomb gas model and transform it into a Euclidean sine-Gordon theory as explained in Sec. II.C. One then performs a Wick's rotation into a quantum sine-Gordon theory in $(1+1)$ dimensions (i.e., one space and one time dimension). This is possible as long as one ignores the precise cutoff prescriptions. The quantum sine-Gordon theory in $(1+1)$ dimensions is, order by order, in perturbation theory equivalent to the $(1+1)$-dimensional massive Thirring model (Coleman, 1975). An alternative route for establishing a fermion connection is to start from a model of a one-dimensional electron gas called the backward scattering model (Menyhard and Solyom, 1973; Luther and Emery, 1974). By using a boson representation of a fermion field (Luther and Emery, 1974; Luther and Peschel, 1974) and factorizing the model into two commuting parts, i.e., a charge-density part and a spin-density part (Luther and Emery, 1974), one can show (Chui and Lee, 1975) that the spin-density part is equivalent to the two-dimensional Coulomb gas model. This connection can also be established without using the boson representation of a fermion (Grinstein *et al.*, 1979). The treatment of the cutoff prescriptions in establishing the fermion-Coulomb gas connections restricts the guaranteed validity of the transformations to small coupling constants (Grinstein *et al.*, 1979). In other words, it is difficult to use the fermion representation for the purpose of obtaining precise results for the Coulomb gas model.

III. RENORMALIZATION EQUATIONS

The renormalization-group equations for the Kosterlitz-Thouless transition describe the critical properties close to the phase transition. In Sec. II.D it was sketched how renormalization-group equations may be constructed using field theory renormalization procedures. The present section will be centered on the physical content of the renormalization-group equations.

As a first step we shall describe in Sec. III.A a more physical, but less systematic, way of obtaining renormalization equations. It may be characterized as a self-consistent linear screening reasoning. The Kosterlitz renormalization-group equations [Eqs. (2.77)] are regained as a lowest-order approximation of this reasoning (Young, 1978; Minnhagen, 1985b). The self-consistent screening approximation may be interpreted in terms of a length-dependent dielectric constant (Kosterlitz and Thouless, 1973; Young, 1978, 1980; Minnhagen, 1985b). This is described in Sec. III.B. In Sec. III.C the properties and implications of the Kosterlitz renormalization-group equations are discussed in some detail. Finally, Sec. III.D discusses some implications of higher-order terms and some possible limitations of the conclusions drawn from the Kosterlitz renormalization-group equations.

A. Linear screening and self-consistent equations

In this section we describe a self-consistent screening reasoning. This reasoning leads to a set of renormalization equations that have a very direct physical interpretation (Minnhagen, 1985b).

The starting point for the reasoning is the charge-density correlation function, which is given by

$$\langle \Delta n(r)\Delta n(0)\rangle = [\langle n^+(r)n^+(0)\rangle + \langle n^-(r)n^-(0)\rangle] - [\langle n^+(r)n^-(0)\rangle + \langle n^-(r)n^+(0)\rangle] \tag{3.1}$$

where $n^{(+)}_{(-)}(r)$ is the density of positive (negative) charges. It turns out that the charge-density correlation function may be expressed as

$$\langle \Delta n(r)\Delta n(0)\rangle = \frac{2z_{\rm eff}^2}{\zeta^2}(e^{-U_{\rm eff}^{++}(r)/T} - e^{U_{\rm eff}^{+-}(r)/T}) , \tag{3.2}$$

where $U_{\rm eff}^{++}(r)=U_{\rm eff}^{--}(r)$ $[U_{\rm eff}^{+-}(r)]$ is the effective interaction between two Coulomb gas charges of equal (opposite) sign [defined so that $U_{\rm eff}(\infty)=0$ for $\lambda_c \neq \infty$] and $z_{\rm eff}$ is the effective particle fugacity, $z_{\rm eff}=e^{\mu_{\rm eff}/T}$, where $\mu_{\rm eff}=T\ln(z)-C_{\rm eff}$ and $C_{\rm eff}$ is a constant. Equation (3.2) is an exact equality, provided $\langle \Delta n(r)\Delta n(0)\rangle$ is the point-charge density (Minnhagen, 1985c). In the sine-Gordon formulation the quantities $U_{\rm eff}^{++}/T$, $U_{\rm eff}^{+-}/T$, and $C_{\rm eff}/T$ turn into a set of sine-Gordon correlation functions.

Equation (3.2) has a simple physical interpretation: $2z_{\rm eff}/\zeta$ may be interpreted as the particle density. Consequently $4z_{\rm eff}^2/\zeta^2$ is the total density of pairs with large separation. Half of the pairs have equal (opposite) signs and contribute to the positive (negative) term of Eq. (3.2). The energy needed to decrease the pair separation from infinity to r of a neutral (non-neutral) pair is given by $-U_{\rm eff}^{+-}(r)$ $(+U_{\rm eff}^{++})$. Consequently the charge-density correlation function has an obvious interpretation as the sum of the effective pair fugacities for pairs with equal and opposite charges, where the particles within a pair are separated by a distance r.

To lowest order in charge, $U_{\rm eff}$ is given by linear-response theory, i.e., $U_{\rm eff}^{+-}(r)=U_{\rm eff}^{++}(r)=V_L(r)$ and similarly $C_{\rm eff}=V_L(0)/2$. To this order Eq. (3.2) reduces to (Minnhagen, 1985b)

$$\langle \Delta n(r)\Delta n(0)\rangle = -\frac{2z^2}{\zeta^2}e^{[V_L(r)-V_L(0)]/T}(1-e^{-2V_L(r)/T}) . \tag{3.3a}$$

Within a linear-response description the linearly screened potential V_L is related to the dielectric function $\varepsilon(r)$ by [compare Eq. (2.35); here, we specialize to the point-charge limit $\hat{f}_{r_0}=1$]

$$\hat{V}_L(k)=\frac{2\pi}{\hat{\varepsilon}(k)k^2} . \tag{3.3b}$$

The dielectric response function $\varepsilon(r)$ is in turn, within a linear-response description, related to the charge-density correlation function through

$$\frac{1}{\hat{\varepsilon}(k)}=1-\frac{2\pi}{Tk^2}\langle \Delta\hat{n}(\mathbf{k})\Delta\hat{n}(-\mathbf{k})\rangle . \tag{3.3c}$$

This means that Eqs. (3.3a)—(3.3c) constitutes a self-consistent set of equations for the linearly screened potential V_L. As described in Sec. II.B.1, the key quantities for characterizing the charge-unbinding transition, i.e., the screening length λ and the dielectric constant $\tilde{\varepsilon}$, may be defined through the linearly screened potential V_L [compare Eq. (2.21a)].

Equations (3.3) are exact in the limit $T\to\infty$, as is apparent because this limit is for constant fugacities equivalent to the small-charge limit $s\to 0$ [compare Eq. (2.10)]. Equations (3.3) are also exact in the limit of large r: for $T>T_c$ this follows since $U_{\rm eff}^{++}$ and $U_{\rm eff}^{+-}$ must reduce to the linearly screened potential V_L for large r because the particle interaction is screened out by free charges. Note that this argument was used in Sec. II.B.1.a in connection with Eq. (2.26) and led to $4T_c\tilde{\varepsilon}(T_c^+)=1$ at the critical line. For $T<T_c$ it follows from the field-theoretic formulation: we want to show that

$$\lim_{r\to\infty}\frac{U_{\rm eff}^{+-}(r)}{V_L(r)}=\lim_{r\to\infty}\frac{U_{\rm eff}^{++}(r)}{V_L(r)}=1 .$$

The quantities $U_{\rm eff}^{+-}/T$, $U_{\rm eff}^{++}/T$, and V_L may be expressed as sine-Gordon correlation functions [compare, for example Eq. (2.61)]. From the field theory formulation it follows that a dimensionless ratio of correlation functions in the limit $r/r_0\to\infty$ can only be a function of $\tilde{\varepsilon}T$ for $T<T_c$ (see, for example, Minnhagen, 1981a). Furthermore, $\tilde{\varepsilon}T$ is constant along a renormalization-group trajectory for $T<T_c$ (compare Sec. III.C). Such a trajectory flows into $z=0$ (compare Sec. III.C). For $z=0$ there is no polarization, consequently $U_{\rm eff}^{+-}=U_{\rm eff}^{++}=V_L$ for $z=0$ and the assertion follows.

The result that $U_{\rm eff}^{+-}$ for large r is given by V_L is to some extent implicit in earlier work. For small T it has been motivated under the assumption that the leading contribution to the polarization comes from dipole pairs with small separations (Nelson, 1978). Close to T_c it is somewhat implicit in the construction of the renormalization-group equations by José *et al.* (1977; see also Nelson, 1983).

One may also note that Eq. (2.24) follows from Eq. (3.3a), since for $T<T_c$ the right-hand side of Eq. (3.3a) is proportional to

$$e^{V_L(r)-V_L(0)}\sim e^{-(1/\varepsilon)\ln(r/r_0)} ,$$

while for $T>T_c$ it is proportional to $V_L(r)\sim e^{-r/\lambda}$ (compare Sec. III.B).

B. Length-dependent screening

We shall introduce the length-dependent screening concept by aid of the quantity $\varepsilon_\infty=\hat{\varepsilon}(k=0)$, where $\hat{\varepsilon}$ is the Fourier transform of the dielectric function $\varepsilon(r)$. The quantity ε_∞ signals the Kosterlitz-Thouless transition in that $\varepsilon_\infty^{-1}\ {=\atop(\neq)}\ 0$ for $T>T_c$ $(T<T_c)$ [see Sec. II.B.1.c and Eq. (2.36)]. For $T<T_c$ the quantity ε_∞ is equal to the dielectric constant $\tilde{\varepsilon}$, whereas for $T>T_c$ it is equal to infinity.

Directly from the definition [Eq. (3.3c)] it follows that

$$1/\varepsilon_\infty = 1+\frac{\pi^2}{T}\int_0^\infty dr\, r^3\langle \Delta n(r)\Delta n(0)\rangle \ . \tag{3.4}$$

This expression lends itself to a simple interpretation in the low-temperature phase. According to the intuitive charge-unbinding picture, the neutral Coulomb gas consists of dipole pairs in this phase (see Sec. II.B.1). The polarizability of a single dipole with separation r, $\alpha(r)$, is given by $\alpha(r)=r^2/2T$. The density of dipoles with intrapair separation r, dip(r) is given by

$$\mathrm{dip}(r) = -\Omega\langle \Delta n(r)\Delta n(0)\rangle/2 \ , \tag{3.5}$$

since only the charges within a dipole pair are correlated. Consequently the electric susceptibility χ_0, corresponding to an "independent dipole pair approximation," is

$$\chi_0 = \frac{1}{\Omega}\int d^2r\,\alpha(r)\mathrm{dip}(r)$$
$$= -\frac{\pi}{2T}\int_0^\infty dr\, r^3\langle \Delta n(r)\Delta n(0)\rangle \ . \tag{3.6}$$

In terms of this quantity ε_∞ is given by

$$\frac{1}{\varepsilon_\infty} = 1-2\pi\chi_0 \ . \tag{3.7}$$

Consequently ε_∞ may be interpreted in terms of polarization due to dipole pairs of all length scales (all sizes of intrapair separations). In electrostatics ε_∞ is related to the electric susceptibility χ by $\varepsilon_\infty = 1+2\pi\chi$, and hence χ is related to the independent pair susceptibility χ_0 by

$$\chi = \frac{\chi_0}{1-2\pi\chi_0} \ . \tag{3.8}$$

One notes that $\chi_0 < 1/2\pi$ in the low-temperature phase and $\chi_0 = 1/2\pi$ in the high-temperature phase.

The length-dependent screening reasoning consists in focusing on how dipoles of different intrapair separations contribute to screening (Kosterlitz and Thouless, 1973). To this end it is expedient to introduce a length-dependent dielectric constant ε_r by

$$\frac{1}{\varepsilon_r} = 1+\frac{\pi^2}{T}\int_0^r dr' r'^3\langle \Delta n(r')\Delta n(0)\rangle \ . \tag{3.9}$$

One notes that ε_r involves all dipoles up to intrapair separation r and that the quantity ε_∞ is given by $\varepsilon_\infty = \varepsilon_{r=\infty}$.

The intrapair force for a dipole with separation r is given by $F=-(d/dr)U_{\mathrm{eff}}^{+-}(r)$ where U_{eff}^{+-} is the effective interaction between a pair of particles with opposite charge. In the absence of surrounding Coulomb gas charges, this force is given by $F_0 = 1/r$. Within linear-response theory (i.e., to lowest order in the charge of the particular dipole under consideration) it is given by $F=-dV_L(r)/dr$, which , by aid of Eqs. (3.3b), (3.3c), and (3.9), may be expressed as (Minnhagen, 1985b)

$$F(r) = F_1(r)+F_2(r)+F_3(r) \ , \tag{3.10a}$$

$$F_1(r) = \frac{1}{r\varepsilon_r} \ , \tag{3.10b}$$

$$F_2(r) = \frac{r\pi^2}{T}\int_r^\infty dr' r'\langle \Delta n(r')\Delta n(0)\rangle \ , \tag{3.10c}$$

$$F_3(r) = \frac{2\pi^2 r}{T}\int_r^\infty dr' r' \ln(r'/r)\langle \Delta n(r')\Delta n(0)\rangle \ . \tag{3.10d}$$

It is interesting to note that these three contributions to the intrapair force, the sum of which constitute the total intrapair force within a linear screening approximation, may be given simple physical interpretations (Minnhagen, 1985b): $F_1(r)$ is caused by the electric field between a dipole pair with intrapair separation r polarized by the dipoles of smaller intrapair separations; $F_2(r)$ is caused by the dipole field outside a pair with separation r polarized by the dipoles of larger separations; while $F_3(r)$ is caused by the orientational energy of a dipole with separation r in the electric field caused by the dipoles of larger intrapair separations.

In the low-temperature phase the screening length λ is infinite and $V_L(r)$ diverges [compare Eq. (2.21a)]. Consequently Eq. (3.3a) reduces to

$$\langle \Delta n(r)\Delta n(0)\rangle = -\frac{2z^2}{\zeta^2}e^{[V_L(r)-V_L(0)]/T} \quad \text{for } T<T_c \ . \tag{3.11}$$

Let us now introduce a logarithmic length scale $l=\ln(r/r_0)$. the effective pair polarizability may then be expressed as [compare Eqs. (3.5) and (3.10)]

$$\alpha(r)\mathrm{dip}(r)d^2r = -\frac{\Omega r^2}{4T}\langle \Delta n(r)\Delta n(0)\rangle d^2r$$
$$= \frac{\Omega\pi}{T}z^2(l)dl \tag{3.12a}$$

with

$$z(l) = z\frac{r_0^2}{\zeta}\exp\left[2l-\frac{1}{2T}\int_0^{r_0e^l} dr' F(r')\right] \ . \tag{3.12b}$$

The quantity $z(l)$, which is defined by Eqs. (3.12), will be referred to as the renormalized fugacity, since this is what it corresponds to in the renormalization-group language (compare Sec. II.D.1). As can be seen from Eq. (3.12a), the physical meaning of $z(l)$ is linked to the polarizability of pairs with the intrapair separation in the interval $[l,l+dl]$. Likewise one may introduce a renormalized temperature $T(l)$ by

$$T(l) = T\varepsilon_{r=r_0\exp(l)} \ . \tag{3.13}$$

in terms of $z(l)$ and $T(l)$, Eqs. (3.3) reduce to a simple set of differential equations in the low-temperature phase (Minnhagen, 1985a, 1985b),

$$\frac{d}{dl}\left[\frac{1}{T(l)}\right]=-\frac{2z^2(l)\pi^2}{T^2}\ , \tag{3.14a}$$

$$\frac{d}{dl}[z(l)]=\frac{z(l)}{2}\left[4-\int_0^\infty dx\, e^{-x}x\Big/T\left[l+\frac{x}{2}\right]\right]\ . \tag{3.14b}$$

The integral in Eq. (3.14b) may by partial integration be expressed as

$$\int_0^\infty dx\, e^{-x}x\Big/T\left[l+\frac{x}{2}\right]=\sum_{n=0}^{\infty}\frac{n+1}{2^n}\frac{d^n}{dl^n}[1/T(l)]$$

$$=1/T(l)+O(z^2(l))\ . \tag{3.15}$$

In a double expansion in the parameters $z(l)$ and $1/T(l)-4$, Eqs. (3.14) reduce to Kosterlitz renormalization-group equations [compare Eq. (2.77)]. Thus to lowest order

$$\frac{d}{dl}\left[\frac{1}{T(l)}\right]=-\frac{2z^2(l)\pi^2}{T^2}\ , \tag{3.16a}$$

$$\frac{d}{dl}[z(l)]=\frac{z(l)}{2}[4-1/T(l)]\ . \tag{3.16b}$$

It may be noted that Eq. (3.16b) follows directly from the definition of $z(l)$ [Eq. (3.12)], provided the intrapair force is estimated by F_1 alone [Eq. (3.10b)], i.e., provided only the polarization due to dipoles with smaller intrapair separations are included (Young, 1978, 1980).

C. Kosterlitz RG equations

The Kosterlitz renormalization-group equations [Eqs. (3.16)] may be explicitly solved. In this section we shall discuss this solution and its implications.

For the purpose of discussing the solution of Kosterlitz RG equations it is convenient to introduce the renormalized temperature variable $t(l)=[1-1/4T(l)]$ and the renormalized fugacity variable $y(l)=\pi z(l)/2T$. In terms of these variables the Kosterlitz RG equations [Eqs. (3.16)] become

$$\frac{d}{dl}[t(l)]=2y^2(l)\ , \tag{3.17a}$$

$$\frac{d}{dl}[y(1)]=2y(l)t(l)\ , \tag{3.17b}$$

which may be directly integrated to

$$t^2(l)-y^2(l)=C\ , \tag{3.18}$$

where C is a constant. Consequently when integrating Eqs. (3.17) from a starting value l_i and the initial conditions $t_i=t(l_i)$ and $y_i=y(l_i)$, one will obtain a $(t(l),y(l))$ point that lies on the trajectory given by Eq. (3.18) with $C=t_i^2-y_i^2$. This is illustrated in Fig. 7. The arrows on the trajectories in the figure indicate the direction in which a $(t(l),y(l))$ point flows with increasing l. There

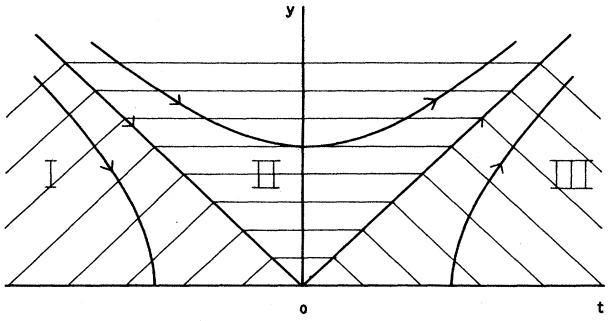

FIG. 7. Flow diagram for the Kosterlitz RG equations [Eqs. (3.17)]. Arrows indicate the flow direction under integration. There are three distinct regions, denoted by I, II, and III, separated by the special trajectories $y=t$ and $y=-t$. The trajectory $y=-t$ corresponds to the critical line (for further explanations see text).

are three distinct regions of initial conditions in the (t,y) plane. Region I is characterized by $C>0$ and $t<0$. A trajectory in this region ends on the t axis at the point $(t(\infty)=-(t_i^2-y_i^2)^{1/2},\ y(\infty)=0)$. Region I is separated from region II by the trajectory $C=0$ and $t(l)=-y(l)$ [compare Eq. (3.18)], which ends on the t axis at the point $(t(\infty)=0, y(\infty)=0)$. Region II is characterized by $y>t$. The trajectories in this region go to (∞,∞). The separation between regions II and III is given by the trajectory $C=0$ and $t(l)=y(l)$, which starts at (0,0). Region III is characterized by $C>0$ and $t>0$. The trajectories in this region start as the t axis and go to (∞,∞).

The explicit solution to Eqs. (3.17) (see, for example, Young and Bohr, 1981) is in region I

$$(t(l),y(l))=(-c\coth(\varphi+2cl),c/\sinh(\varphi+2cl))\ ,$$

$$\cosh(\varphi)=-t_i/y_i\ ; \tag{3.19a}$$

in region II

$$(t(l),y(l))=(c\cot(\varphi-2cl),c/\sin(\varphi-2cl))\ ,$$

$$\cos(\varphi)=t_i/y_i,\quad 0<\varphi<\pi\ ; \tag{3.19b}$$

in region III

$$(t(l),y(l))=(c\coth(\varphi-2cl),c/\sinh(\varphi-2cl))\ ,$$

$$\cosh(\varphi)=t_i/y_i\ ; \tag{3.19c}$$

and $c=(\,|t_i^2-y_i^2|\,)^{1/2}$ in all three cases.

In region I, ε_∞ is obtained from $(t(\infty)=-(t_i^2-y_i^2)^{1/2}, y(\infty))$, i.e.,

$$\varepsilon_\infty=[4T+4T_c[\sqrt{(1-T/T_c)(\pi z_i/T_c+1-T/T_c)}]^{-1} \tag{3.20a}$$

with

$$T_c=\frac{1}{4\varepsilon_i}-\frac{\pi z_i}{2} \tag{3.20b}$$

and $\varepsilon_i=\varepsilon(l_i)$. The separation trajectory between regions I and II corresponds to $T=T_c$ and to the critical line (as will be motivated further below). Consequently, as the critical line is approached from below, ε_∞ behaves as

$$\varepsilon_\infty=\varepsilon_c(1-c_<\sqrt{1-T/T_c})+O(1-T/T_c) \tag{3.21}$$

with $\varepsilon_c=1/4T_c$ and $c_<=2\sqrt{\pi z_i\varepsilon_i}$. This is in accord with the results of Sec. II.B.1 [compare Eqs. (2.23), (2.30), and (2.32)].

In regions II and III it is not meaningful to integrate Eqs. (3.17) to infinity since $y(l)$ divergence at $l=l_0=\varphi/2c$. Furthermore, the guaranteed validity of the equations is restricted to $y(l)\ll 1$. Nevertheless, it is possible to deduce how the screening length λ diverges as the critical line is approached (Kosterlitz, 1974). The following argumentation is due to Young and Bohr (1981).

First one observes that integrating the equations is the same as mapping a Coulomb gas specified by (t_i,y_i,r_0) onto one specified by $(t(l),y(l),r_0e^l)$. All such Coulomb gases have identical critical properties obtained by integrating to $l=\infty$ (Wilson, 1971). This means that the dimensionless quantity $\tilde{\lambda}=\lambda/r_0e^l$ can only be a function of the two dimensionless variables $(t(l),y(l))$ and that

$$\frac{\lambda}{r_0}=e^l\tilde{\lambda}(t(l),y(l)) \tag{3.22}$$

for all $(t(l),y(l))$ points along a trajectory, since λ governs the critical behavior. Equation (3.22) is just the basic statement of how a correlation length scales under a renormalization-group transformation (Wilson, 1971).

The idea is now to combine Eqs. (3.19) with Eq. (3.22). First we impose the restriction $2c(l_0-l)=\varphi-2cl\ll 1$, which means that $t(l)\simeq y(l)$ [see Eqs. (3.19)] or, in other words, we only consider trajectories that, for some large enough l, come close to the line $t=y$. In addition we impose the condition $l_0-l\gg 1$. From Eqs. (3.19) it then follows that

$$t(l)\simeq y(l)\simeq\frac{1}{2(l_0-l)}\ll 1\ , \tag{3.23}$$

with strict equalities on the line $t=y$. Consequently the additional imposed condition ensures that we are still inside the expected validity range of Eqs. (3.19). From Eq. (3.23) it follows that $l\simeq l_0-1/[2y(l)]$ with strict equality on the special trajectory $t(l)=y(l)$. Inserting this result into Eq. (3.22) gives

$$\frac{\lambda}{r_0}e^{-l_0}\simeq e^{-1/2y(l)}\tilde{\lambda}(t(l),y(l)) \tag{3.24}$$

with strict equality on the trajectory $t(l)=y(l)$. Consequently, on this trajectory, one has

$$e^{-1/2y(l)}\tilde{\lambda}(t(l),y(l))=\bar{C}\ , \tag{3.25}$$

where, since the left-hand side of Eq. (3.24) is independent of l, $\bar{C}$ is a fixed positive constant. On physical grounds one expects that $e^{-1/2y}\lambda(t,y)$ is finite and analytic in the parameter region $t>0$ and $y>0$, since the screening length λ is finite in this region and no additional phase transition is expected. Consequently, for any trajectory that comes close enough to the line $t=y$, the leading contribution to the screening length is given by (Young and Bohr, 1981)

$$\frac{\lambda}{r_0}=e^{l_0}\bar{C}\ . \tag{3.26}$$

From Eq. (3.18) it follows that the closer a point is to the critical line $t=-y$ in regions II and III, the closer the corresponding trajectory comes to the line $t=y$. Consequently, as the critical line is approached, the dominant contribution to the screening length is given by Eq. (3.26).

Let us first consider the case in which a point $y\neq 0$ on the critical line is approached. Equation (3.19b) applied to this case gives

$$l_0=\frac{\varphi}{2c}\simeq\frac{\pi}{2y}\frac{1}{(1-u^2)^{1/2}}-\frac{1}{2y}\quad\text{for } u=\frac{t}{y}\to -1\ , \tag{3.27}$$

which translated back to the (T,z) variables becomes (Young and Bohr, 1981)

$$\lambda\simeq\bar{C}\exp\left[\tfrac{1}{2}\sqrt{\pi T_c/z_i}\frac{1}{\sqrt{T/T_c-1}}\right]e^{-T_c/(2\pi z_i)} \tag{3.28}$$

for $T\to T_c$. Thus the screening length diverges as the critical line is approached from above. This, in turn, may be taken as an *ipso facto* verification that the $t=y$ line is correctly identified as the critical line. One notes that the form of the divergence (first obtained by Kosterlitz, 1974) agrees with the result of Sec. II.B.1, which was inferred from a hand-waving argument for the behavior of the dielectric constant $\tilde{\varepsilon}(T)$ close to T_c [compare Eqs. (2.32)–(2.34)].

As a second case let us consider the approach to the $(t=0,y=0)$ point on the critical line following a path in the (t,y) plane for which $y\ll t$. Equation (3.19c) applied to this case gives

$$l_0=\frac{\varphi}{2c}\simeq\frac{\ln(2u)}{2y(u^2-1)^{1/2}}\quad\text{for } u=\frac{t}{y}\to\infty\ , \tag{3.29}$$

which translated back to the (T,z) variables becomes

$$\lambda\simeq\bar{C}\left[\frac{\pi z}{2T-1}\right]^{-2T/(4T-1)}\ , \tag{3.30}$$

in agreement with the result of Sec. II.B.1.a [compare Eq. (2.30)] and the sine-Gordon result (compare Sec. II.D.1 and Eq. (2.71)].

More generally one obtains from Eqs. (3.19) that l_0 may be expressed as $l_0=f(u)/2y$, with $u=t/y$ and $f(u)$ a continuous function in the interval $-1<u<\infty$. Hence the dominant contribution to λ is given by

$$\lambda=\bar{C}e^{f(u)/2y},\quad u=t/y \tag{3.31}$$

as the critical line is approached (Young and Bohr, 1981). The function $f(u)$ is plotted in Fig. 8.

The singular part of the free energy F_s (i.e., the part of the free energy that contains the nonanalytic behavior caused by the phase transition) can be obtained along the

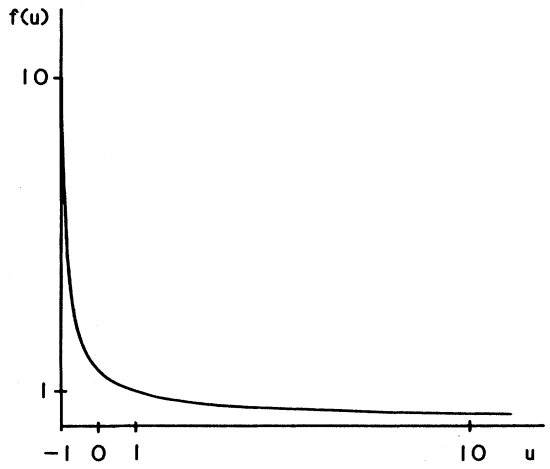

FIG. 8. Plot of the function $f(u)$ appearing in the expression for the screening length λ [Eq. (3.31)].

$$f(u)=\begin{cases}\dfrac{\arccos(u)}{\sqrt{1-u^2}} & \text{for } -1<u\le 1,\\ \dfrac{\ln(u+\sqrt{u^2-1})}{\sqrt{u^2-1}} & \text{for } u\ge 1.\end{cases}$$

same lines as the screening length (Young and Bohr, 1981). In analogy with Eq. (3.22) the dimensionless quantity $\tilde{F}_s=r_0^2e^{2l}F_s$ can only be a function of the dimensionless variables $(t(l),y(l))$ and is constant along a RG trajectory (Wilson, 1971). Combining this with Eqs. (3.19) and repeating the argument leading to Eq. (3.26) gives $F_s\sim e^{-2l_0}$, which by use of Eq. (3.26) becomes

$$F_s\sim\lambda^{-2}\ . \tag{3.32}$$

This is in agreement with the result for F in Sec. II.B.1.d [compare Eqs. (2.30a) and (2.38a)]. It means that the singular part F_s may be interpreted as the free energy caused by unbound Coulomb gas particles (denoted by F_F in Sec. II.B.1.d).

It is also possible to relate the regular part of the free energy F_B [see Eq. (2.39a) and (2.72c)] to Kosterlitz RG equations (Young and Bohr, 1981). To this end one may use the exact relation (Minnhagen, 1985c)

$$z^2\frac{\partial^2F}{\partial\bar{z}_+\partial\bar{z}_-}\bigg|_{\Delta E=0}=-\frac{2Tz_{\rm eff}^2}{\zeta^2}\int d^2r(e^{U^{+-}(r)/T}-1)\ , \tag{3.33}$$

where $\bar{z}_{\substack{+\\(-)}}$ was introduced in Eq. (2.56). U^{+-} and $z_{\rm eff}$ are the same as in Eq. (3.2). Starting from Eq. (3.33) we may in analogy with Eqs. (3.3a) approximate F_B for small z [where $z(l)$ is defined by Eq. (3.12b)] by

$$F_B=z^2\frac{\partial^2F}{\partial\bar{z}_+\partial\bar{z}_-}\bigg|_{\Delta E=0}$$
$$=-\frac{4\pi T}{r_0^2}\int_0^\infty dl\,e^{-2l}z^2(l)(1-e^{-V_L(r_0e^l)/T})\ , \tag{3.34a}$$

which below T_c, since $V_L=\infty$ for $T<T_c$, reduces to

$$F_B=-\frac{4\pi T}{r_0^2}\int_0^\infty dl\,e^{-2l}z^2(l)\ . \tag{3.34b}$$

Suppose that we apply an *ad infinitum* repeated partial integration to Eq. (3.34b) such that the e^{-2l} part of the integrand is integrated in each step. In addition we assume that a set of RG equations exists and that they are of the form

$$\left[\frac{d}{dl}z(l)=\beta_z(t(l),z(l)),\frac{d}{dl}t(l)=\beta_t(t(l),z(l))\right],$$

where the β functions can be expanded in powers of $t(l)$ and $z(l)$. Note that the $l=\infty$ limit of the partial integration does not contribute, since $(z(\infty),t(\infty))=(0,1-1/4T\tilde{\epsilon})$ for $T<T_c$. It then follows that

$$F_B=-\frac{4\pi T}{r_0^2}K(t(0),z(0))\ , \tag{3.35}$$

where K is a function that can be expanded in powers of $t(0)$ and $z(0)$. By this procedure one can calculate K to leading order in z using Kosterlitz RG equations [Eqs. (3.17); Young and Bohr, 1981],

$$F_B=\frac{4\pi T^2z^2}{\zeta(2T-1)}\ , \tag{3.36}$$

in perfect agreement with Eq. (2.72c). In the same way the coefficients in front of the higher-order z terms obtained by using the Kosterlitz RG equations correctly reproduce the singularities at the temperatures $T_p=p/2(2p-1)$ [see Sec. II.B.1.d below Eq. (2.72c); Young and Bohr, 1981]. Note, however, that the argument given above presumed $T<T_c$ [whereas Eqs. (2.72) were constructed for $T>T_c$]. The above approach may, however, be extended to $T>T_c$ (Young and Bohr, 1981) without any change in the result for the regular part of the free energy F_B, i.e., the part of the free energy that may be associated with bound pairs [compare Eqs. (2.39)].

This ends our survey of the solution and implications of Kosterlitz renormalization-group equations. In the next section we shall discuss some possible limitations and extensions.

D. Limitations and extensions of Kosterlitz RG equations

One systematic way of improving Kosterlitz renormalization-group equations is to use field-theoretic renormalization-group procedures (see, for example, Amit, 1984). The field theory corresponding to the

Coulomb gas model is a sine-Gordon theory (compare Sec. II.C). The systematic procedure presumes the existence of renormalization-group equations of the form

$$\left[\frac{d}{dl}(t_R(l))=\beta_t(t_R(l),z_R(l)) \ , \ \frac{d}{dl}(z_R(l))=\beta_z(t_R(l),z_R(l))\right]$$

where the β functions at the critical line can be obtained, order by order, in double expansions in the variables (t_R,z_R) by keeping track of the logarithmic singularities in a diagrammatic expansion (see Sec. II.D.1). This procedure has been carried out to one order higher than have the Kosterlitz RG equations by Amit *et al.* (1980), resulting in Eqs. (2.78). The difference from a practical point of view between these equations and the Kosterlitz RG equations is illustrated in Fig. 9, where the critical lines obtained from the respective equations are plotted. In this comparison, and in others in this section, we use for simplicity the initial condition $T_R(l=0)=T(l=0)=T$, $z_R(l=0)=z(l=0)=z$.

One notes that the two sets of RG equations given by Eqs. (3.16) and (2.78), respectively, both suggest that the critical line is identical to the trajectory defined by $T\varepsilon_\infty(T)=\frac{1}{4}$. However, one may also note from the derivation of Eqs. (2.78) that the β functions are obtained as double expansions in the variables $(T_R-\frac{1}{4},z_R/T)$ and that, consequently, the expansions can only *a priori* be trusted very close to $T=\frac{1}{4}$ and $z=0$.

Nevertheless one may argue that the critical properties obtained from the Kosterlitz RG equations (see Sec. III.C) have a wider validity in the following way: *If* a part of the critical line consists of a RG trajectory that flows towards small z, *then* this trajectory will eventually enter the region where the expansions of the β functions are valid and hence eventually the region where the Kosterlitz RG equations are valid. Consequently the critical properties derived from Kosterlitz RG equations should be valid for that part of the critical line which consists of such a RG trajectory. For example, $T\varepsilon_\infty(T)=\frac{1}{4}$ for all points on such a part of the critical line. This argument constitutes the basis for the universal jump prediction [see Sec. II.B.1.c, Eq. (2.36), and Sec. IV.C].

However, it should be observed that the above argument leaves open the question of where such a part of the critical line would start. Figure 9 suggests that it should start at $T=0,z=$finite, in which case the whole critical line would consist of the trajectory ending at $T=\frac{1}{4}$. Another possibility is that only a part of the critical line close to $T=\frac{1}{4}$ consists of this trajectory, which in such a case must start at some specific point in the (T,z) plane (Minnhagen, 1985a, 1985b).

Another possibility for improving the Kosterlitz RG equations is provided by the self-consistent linear screening reasoning described in Sec. III.A. This reasoning leads to the self-consistent equations (3.3). In the low-temperature phase these equations reduce to the renormalization equations (3.14) (Minnhagen, 1985a, 1985b),

$$\frac{d}{dl}\left[\frac{1}{T(l)}\right]=-\frac{2z^2(l)\pi^2}{T^2} \ , \tag{3.14a}$$

$$\frac{d}{dl}[z(l)]=\frac{z(l)}{2}\left[4-\int_0^\infty dx\, e^{-x}x/T(l+x/2)\right] . \tag{3.14b}$$

In Fig. 10 the critical line obtained from Eqs. (3.14) is compared to those obtained from Kosterlitz RG equations [Eqs. (3.16)] and the next-order RG equations [Eqs. (2.78)]. One notes that Eqs. (3.14) and (2.78) agree quite well in that they give very similar corrections to the critical line obtained from Kosterlitz RG equations [Eqs. (3.16)].

Figure 11 shows the flow diagram for $(T(l),z(l))$

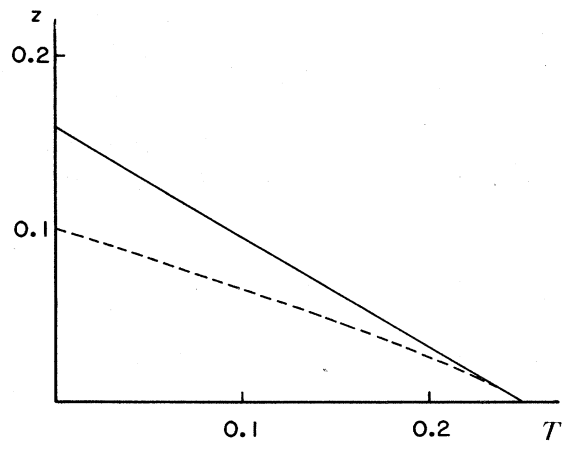

FIG. 9. The Kosterlitz-Thouless phase-transition line in the (T,z) plane as obtained from the two lowest-order RG equations: solid line, lowest-order RG equations [Eqs. (3.16)]; dashed curve, next-order RG equations [Eqs. (2.78)].

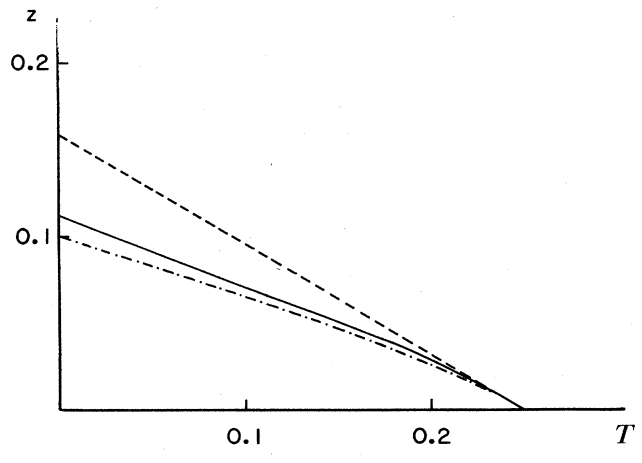

FIG. 10. The Kosterlitz-Thouless phase-transition line in the (T,z) plane obtained by three different approximations: solid curve, the self-consistent linear screening approximation [Eqs. (3.14)]; dashed line, lowest-order RG equations [Eqs. (3.16)]; dot-dashed curve, next-order RG equations [Eqs. (2.78)].

points in the (T,z) plane as obtained from Eqs. (3.14) (Minnhagen, 1985a, 1985b). The most striking feature is that the trajectory ending at the point $(\frac{1}{4},0)$ starts at a specific point (T^*,z^*) $[\simeq(0.144,0.054)]$ in the (T,z) plane. Consequently this trajectory constitutes only one part of the critical line. The other part of the critical line, i.e., the critical line for temperatures lower than T^* (dashed line in Fig. 11) is determined by the line consisting of the starting points for all trajectories ending on the T axis.

In Fig. 12 the quantity $1/(T_c\varepsilon_c)$ is plotted as a function of T_c where $\varepsilon_c=\lim_{T\to T_c^-}\varepsilon_\infty(T)$. In the interval $T^*<T_c<\frac{1}{4}$ the value of this quantity is 4, which corresponds to the universal jump value (compare Table I and Sec. IV.C). For $T_c<T^*$ the value obtained from Eqs. (3.14) is larger and nonuniversal. The behavior of $\varepsilon_\infty(T)$ as the critical line is approached from below for constant z is

$$\varepsilon_\infty(T)=\varepsilon(T_c^-)-\text{const}\times\sqrt{T_c-T}\ , \tag{3.37}$$

in accordance with Eq. (2.32) of Sec. II.B.1.b.

The important point to note is that Eqs. (3.14) imply that the universal jump prediction breaks down for a two-dimensional Coulomb gas with a critical temperatures lower than the specific temperature T^*.

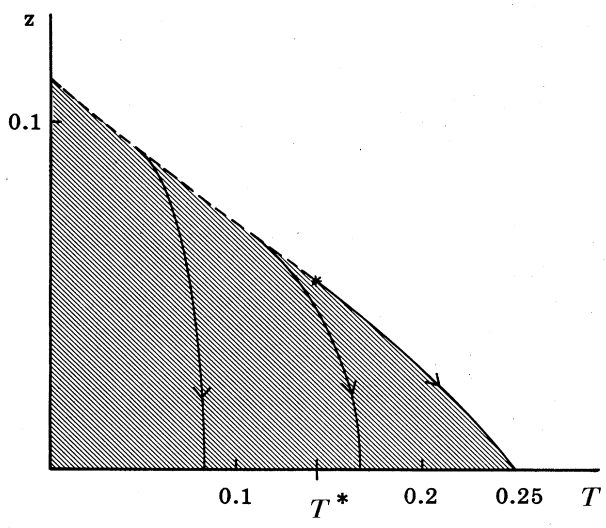

FIG. 11. Flow diagram for $(T(l),z(l))$ points in the low-temperature phase corresponding to the self-consistent linear screening approximation given by Eqs. (3.14): shaded area, low-temperature phase; solid curves with arrows, $(T(l),z(l))$ trajectories with constant $\tilde{\varepsilon}T$ (the arrows indicate the direction of increasing l); dashed curve, the loci of starting points for trajectories with constant $\tilde{\varepsilon}T$. The starting point (T^*,z^*) of the special trajectory $\tilde{\varepsilon}T=\frac{1}{4}$ is denoted by an asterisk in the figure. The trajectory $\tilde{\varepsilon}T=\frac{1}{4}$ constitutes the critical line for $T^*\le T_c\le\frac{1}{4}$. For $T_c<T^*$ the critical line is given by the dashed curve, i.e., by the starting points of the trajectories in the low-temperature phase.

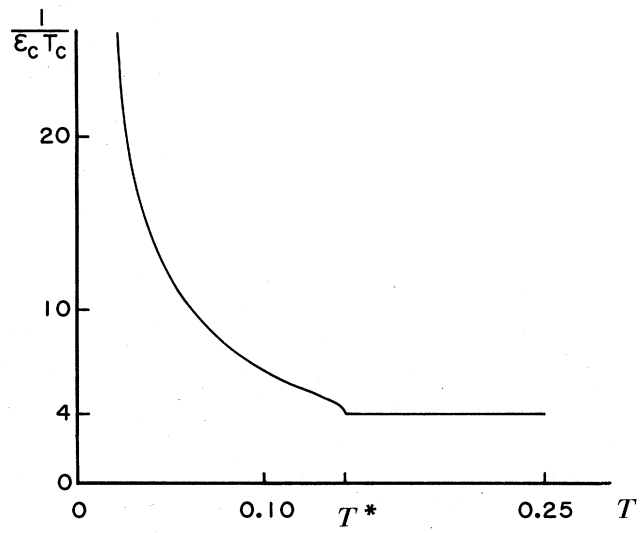

FIG. 12. The jump of the quantity $1/[\hat{\varepsilon}(k=0)T]$ at the critical line as obtained from the self-consistent linear screening approximation [Eqs. (3.14)],

$$1/[\hat{\varepsilon}(k=0)T]=\begin{cases}1/(T_c\varepsilon_c) & \text{for } T\to T_c^-\ ,\\ 0 & \text{for } T\to T_c^+\ .\end{cases}$$

The quantity $1/(T_c\varepsilon_c)$ is plotted. In the interval $T^*\le T_c\le\frac{1}{4}$ the value is given by the universal jump prediction $1/(\varepsilon_c T_c)=4$. Below T^* the value of the jump is larger and nonuniversal.

What then could be the physics behind the implied change of behavior at the point (T^*,z^*) on the critical line? One possibility is the following: In the intuitive charge-unbinding picture, the Kosterlitz-Thouless transition occurs when the first pair unbinds. This suggests two distinct possibilities—either the phase transition involves an infinitesimal fraction of the pairs or a finite fraction. One possible scenario is then that the high-temperature part of the critical line corresponds to the first possibility, while the low-temperature part corresponds to the second (Minnhagen and Wallin, 1987). This would further suggest that the low-temperature part of the transition is of first order (Caillol and Levesque, 1986).

This ends our survey of some possible extensions of the Kosterlitz RG equations.

IV. TWO-DIMENSIONAL SUPERFLUIDS AND VORTICES

In the preceding sections the properties of the two-dimensional Coulomb gas have been reviewed in some detail. Here we outline how the two-dimensional Coulomb gas provides a description of vortex fluctuations for two-dimensional superfluids.

This subject derives some of its general interest from the following facts: First, the superfluid state of a two-dimensional superfluid is destroyed by thermally excited

vortices (Berezinskii, 1971; Kosterlitz and Thouless, 1972). This means that thermally excited vortices are strongly reflected in the behavior of a two-dimensional superfluid for temperatures close to the superfluid transition. Second, there exist physical realizations of two-dimensional superfluids accessible to high-precision experiments. Examples are, in the case of a neutral two-dimensional superfluid, ^{4}He films on suitable substrates (see, for example, Bishop and Reppy, 1978, who designed the first experiment to test the Kosterlitz-Thouless aspect of the superfluid transition) and, in the case of a charged superfluid, thin superconducting films (see, for examples, Beasley *et al.*, 1979, who were the first to point out the experimental possibility of studying the Kosterlitz-Thouless transition on superconducting films).

The focus of the presentation will be on the effects of thermally excited vortices close to the superfluid transition in two dimensions and on how these effects are reflected in actual experiments. The presentation will be rather phenomenological.

Our starting point is an estimate of the energy governing the thermodynamic fluctuations of a two-dimensional superfluid. For a neutral superfluid, this is taken to be the kinetic energy due to the superfluid mass flow (Sec. IV.A.1). In the case of a charged superfluid, this kinetic energy has to be supplemented with modifications due to the coupling to the electromagnetic field (Sec. IV.B.1). At finite temperatures the fluctuating superfluid mass flow will spontaneously form vortex configurations. A description of these thermal vortex fluctuations leads to a model that is isomorphic to the two-dimensional Coulomb gas. This model and the mapping to the Coulomb gas are spelled out in some detail for ^{4}He films (Sec. IV.A.2) and for superconductors (Sec. IV.B.2). The point is that the results for the two-dimensional Coulomb gas (as described in Secs. II and III) via these mappings have direct bearing on the superfluid-superconducting films. An important example is the Coulomb gas prediction for the superfluid density (Sec. IV.C.).

As a next step the Ginzburg-Landau Coulomb gas model is introduced (Sec. IV.D.1). This phenomenological model leads to the concept of Coulomb gas scaling (Minnhagen, 1981b). The Coulomb gas scaling relations offer a particularly direct bridge between theory and experiment (Sec. IV.D.2).

In the case of a regular two-dimensional array of superconducting grains, a natural starting description of the energy governing the thermodynamic fluctuations focuses on the nearest-neighbor order-parameter phase coupling between the grains. This leads to XY-type models. The relation between this type of lattice model and a continuum description in terms of the superfluid mass flow is taken up in Sec. IV.E.

Explicit examples of experimental results reflecting "static" Coulomb gas properties (as opposed to experimental results reflecting the dynamical aspects of the vortices) will be given in Sec. V. Modifications due to the dynamical aspects of the vortices will be discussed in Sec. VI.

A. Neutral superfluid

We now turn to a simple phenomenological description of the thermodynamic fluctuations for a neutral superfluid.

It is assumed that the superfluid state at each point $\mathbf{r}$ may be characterized by a complex order parameter $\psi(\mathbf{r})=\sqrt{\rho(\mathbf{r})}e^{i\theta(\mathbf{r})}$, where $\rho(\mathbf{r})$ is the superfluid mass density ("density" implies "areal density" in the present two-dimensional context) and the phase θ is related to the superfluid velocity v_S (see, for example, Baym, 1969; Forster, 1975) by

$$\mathbf{v}_S(\mathbf{r})=\frac{\hbar}{m^*}\nabla\theta(\mathbf{r}) . \tag{4.1}$$

Here $\hbar$ is Planck's constant divided by 2π, and m^* is the mass of the "boson" causing the superfluidity, i.e., for ^{4}He it is the mass of the helium atom and in the case of a superconductor it is twice the electron mass (=the mass of a Cooper pair) (see, for example, Baym, 1969).

A two-dimensional superfluid, from the point of view of a "real" three-dimensional superfluid layer, implies that the thickness of the layer is small enough that we may neglect any superflow in the direction perpendicular to the layer (see Fig. 13). A criterion is given in connection with the Ginzburg-Landau description in Sec. IV.D.1. Physical quantities like, for example, the mass density $\rho(\mathbf{r})$ can, from this point of view, be interpreted as averages over the layer thickness, i.e.,

$$\rho(\mathbf{r})=\int_{-d/2}^{d/2}dx_3\rho(\mathbf{r},x_3) ,$$

where $\rho(\mathbf{r},x_3)$ is the three-dimensional mass density (see Fig. 13).

1. Configuration energy

We begin with the assumption that thermodynamic properties are determined by fluctuations in the superfluid mass flow and that the energy for these fluctuations is the kinetic energy associated with the mass flow. This kinetic energy is taken to be

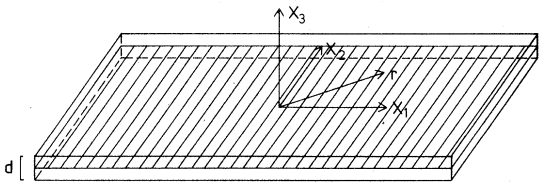

FIG. 13. Sketch of a superfluid layer. The $\hat{\mathbf{x}}_3$ direction is perpendicular to the layer. The thickness of the layer is d. For a thin enough layer the superflow in the $\hat{\mathbf{x}}_3$ direction can be neglected and any dependence on the x_3 coordinate suppressed. This reduces the description to that of a two-dimensional superfluid sheet (shadowed $x_3=0$ plane in the figure). The position vector on this sheet is denoted by $\mathbf{r}=(x_1,x_2)$.

$$H_S=\frac{1}{2\rho}\int d^2r\,\mathbf{g}^2(\mathbf{r})\ , \tag{4.2}$$

where $\mathbf{g}(\mathbf{r})$ is the mass-current-density field and ρ is the mass density of the homogeneous superfluid in the absence of any superfluid mass flow.

In principle one may imagine that a description in terms of a "smooth" mass-current field, as is implied by Eq. (4.2), originates from a microscopic description via a "coarse-graining" process in which all degrees of freedom up to some suitable length scale have been taken into account by some renormalization procedure. The reader is referred to the article by Halperin (1979) for a discussion of this point. The attitude taken in this article is that the H_S given by Eq. (4.2) constitutes a plausible phenomenological starting point. However, it is conceptually important to note that both the quantity ρ (the "bare" superfluid density) and the quantity $\mathbf{g}$ (the superfluid mass-current-density field) used to specify the configuration energy H_S given by Eq. (4.2) are length-scale dependent. Thus in order for H_S to be completely interpretable it is in principle necessary to specify in addition the length scale involved in defining ρ and $\mathbf{g}(\mathbf{r})$. The fluctuations on smaller length scales are then absorbed in ρ, which consequently in general depends on the temperature (Halperin, 1979).

The thermodynamical averages are obtained via the partition function $Z=\mathrm{Tr}e^{-H_S/T}$ where T is the temperature (in units such that the Boltzmann factor is suppressed). The trace implies a functional integration over all possible mass-current fields $\mathbf{g}(\mathbf{r})$.

The mass current is conveniently separated into longitudinal and transverse parts,

$$\mathbf{g}(\mathbf{r})=\mathbf{g}_{||}(\mathbf{r})+\mathbf{g}_{\perp}(\mathbf{r})\ , \tag{4.3a}$$

which have zero curl and divergence, respectively,

$$\nabla\times\mathbf{g}_{||}(\mathbf{r})=0\ , \tag{4.3b}$$

$$\nabla\cdot\mathbf{g}_{\perp}(\mathbf{r})=0\ . \tag{4.3c}$$

Accordingly these mass currents can be written in terms of potentials. The longitudinal part is the gradient of a scalar field Φ,

$$\mathbf{g}_{||}(\mathbf{r})=\nabla\Phi(\mathbf{r})\ , \tag{4.4a}$$

while the transverse part may be expressed in a scalar field W as

$$\mathbf{g}_{\perp}(\mathbf{r})=\nabla\times\hat{\mathbf{x}}_3W(\mathbf{r})\ , \tag{4.4b}$$

where $\hat{\mathbf{x}}_3$ is the unit vector perpendicular to the plane of the two-dimensional superconductor and the $\mathbf{r}$ vector is in the plane of the superfluid (see Fig. 13).

The field $W(\mathbf{r})$ is restricted by the fact that the state of a superfluid is expressible by a complex order parameter $\psi(\mathbf{r})=\sqrt{\rho(\mathbf{r})}e^{i\theta(\mathbf{r})}$. This restriction may be approximately taken into account in the following way: The single-valuedness of ψ requires that any line integral around a closed loop of the gradient of the phase $\theta(\mathbf{r})$ be an even multiple of π,

$$\int d\mathbf{s}\cdot\nabla\theta(\mathbf{r})=2s\pi\ , \tag{4.5}$$

where s is an integer. Obviously, a nonzero s is consistent with the single-valuedness of ψ and the continuity of $\rho(\mathbf{r})$ and $\theta(\mathbf{r})$ *only if* ψ vanishes at one or more points inside the closed loop. An isolated point where ψ vanishes may hence be characterized by an integer s, where s is defined as the value divided by 2π of the line integral of $\nabla\theta$ around any closed loop surrounding this (and only this) particular $\psi=0$ point. By convention the direction of the line integral is taken to be counterclockwise with respect to the $\hat{\mathbf{x}}_3$ axis.

An isolated $\psi=0$ point with $s\neq0$ describes a vortex in the superfluid, and s is the corresponding vorticity. This follows since the superfluid velocity v_S is proportional to $\nabla\theta$ [see Eq. (4.1)]. Consequently, if $\int d\mathbf{s}\cdot\nabla\theta\neq0$ for a loop around a $\psi=0$ point, then the superfluid mass flow is circulating around this point in correspondence with the usual definition of a vortex in a fluid.

The quantized vorticity for a superfluid imposes restrictions on the superfluid mass flow. Consider a single vortex with vorticity s at $\mathbf{r}=0$. The mass-current density $\mathbf{g}$ is related to the superfluid velocity field by $\mathbf{g}=\rho(\mathbf{r})\mathbf{v}_S(\mathbf{r})$. *If* $\rho(\mathbf{r})$ can be approximated by a constant, *then* Eq. (4.5) may, by use of the Stokes theorem, be directly transformed into a condition for the scalar field W describing the transverse part of the mass-current density [compare Eqs. (4.1) and (4.4b)],

$$\nabla^2W(\mathbf{r})=-2\pi s\frac{\rho\hbar}{m^*}\delta(\mathbf{r})\ . \tag{4.6}$$

However, the δ function on the right-hand side implies an infinite mass flow at the vortex center, which is obviously unphysical. The fact that the mass flow has to be finite everywhere means that $\rho(\mathbf{r})$ has to vanish at the vortex center, which just reflects the physical significance of the condition $\psi=0$ at the vortex center. This variation of $\rho(\mathbf{r})$ close to the vortex center may be approximately taken into account by modifying Eq. (4.6) into

$$\nabla^2W(\mathbf{r})=-2\pi s\frac{\rho\hbar}{m^*}f_{r_0}(\mathbf{r})\ , \tag{4.7}$$

where f_{r_0} is a function that is everywhere finite and that approaches zero for $r>r_0$ and is normalized, so that $\int d^2r\,f_{r_0}(\mathbf{r})=1$. The region $r<r_0$ is termed the vortex core region; it is the region where the deviation of $\rho(\mathbf{r})$ from a constant value becomes significant. The functions f_{r_0} (one for each s) and the relation between f_{r_0} and $\rho(\mathbf{r})$ can be obtained within a Ginzburg-Landau approximation (Minnhagen and Nylén, 1985a; see Sec. VI.D.1).

Thus, in short, the quantized vorticity for a superfluid imposes restrictions on the scalar field W describing the transverse part of the superfluid mass flow. This restriction is approximately expressed by Eq. (4.7). Consequently, an arbitrary transverse mass flow can be specified by giving the positions of the corresponding vortices together with the respective vorticities and "source" functions f_{r_0}.

The longitudinal part of the mass flow can be described by a scalar field Φ [Eq. (4.4a)]. A convenient choice of boundary condition is that of Φ constant on the walls of the superfluid, so that there is no net flow across the walls. A constant net flow may then be imposed by adding a constant vector to $\nabla\Phi$ [compare Eq. (4.4a)]. With this boundary condition H_S reduces to

$$H_S=\frac{1}{2\rho}\int d^2r\,\mathbf{g}^2(\mathbf{r}) =\frac{1}{2\rho}\int d^2r[\mathbf{g}_{||}^2(\mathbf{r})+\mathbf{g}_{\perp}^2(\mathbf{r})]\,, \tag{4.8}$$

since the cross term $\int d^2r\,\mathbf{g}_{||}\cdot\mathbf{g}_{\perp}$ vanishes by partial integration. Thus the longitudinal and transverse parts of the flow separate completely in the absence of a net flow across the walls.

We are aiming at a description of how the superfluidity is destroyed and how this is reflected in various macroscopic quantities. A simple way of defining the superfluid property within the present description is to consider the total increase in free energy ΔF when a total infinitesimal constant net flow $\delta\mathbf{G}_{\rm tot}$ is imposed across the sample. This kinetic energy must be proportional to $(\delta\mathbf{G}_{\rm tot})^2$, and it is intuitively clear that the constant of proportionality measures the macroscopic superfluid density ρ_S, i.e.,

$$\Delta F=\frac{1}{2\Omega\rho_S}(\delta\mathbf{G}_{\rm tot})^2 \tag{4.9a}$$

or equivalently

$$\rho_S^{-1}=\Omega\left[\frac{\partial^2 F}{\partial G_{\rm tot}^2}\right]_{G_{\rm tot}\to 0}, \tag{4.9b}$$

where Ω is the (two-dimensional) volume. Now, since $F=T\ln Z=-T\ln(\mathrm{Tr}e^{-H_S/T})$, a straightforward calculation starting from Eqs. (4.2) and (4.9b) gives ρ_S in terms of correlation functions as (Minnhagen and Warren, 1981)

$$\rho_S=\frac{1}{T}\int d^2r[\langle\mathbf{g}_{||}(\mathbf{r})\cdot\mathbf{g}_{||}(0)\rangle-\langle\mathbf{g}_{\perp}(\mathbf{r})\cdot\mathbf{g}_{\perp}(0)\rangle]\,, \tag{4.10}$$

where $\langle\ \rangle$ denotes the thermodynamical average with respect to the configuration energy H_S. This definition in terms of correlation functions, as well as more discussion on how to define the macroscopic superfluid density, can be found, for example, in Hohenberg and Martin (1965), Baym (1969), and Forster (1975).

The first correlation function on the right-hand side of Eq. (4.10) involves only the longitudinal part of the mass-current field, since the configuration energy H_S decouples into longitudinal and transverse parts [see Eq. (4.8)]. The thermal average hence reduces to a functional integral over $\mathbf{g}_{||}$, which can be carried out directly,

$$\int d^2r\langle\mathbf{g}_{||}(\mathbf{r})\cdot\mathbf{g}_{||}(0)\rangle=\rho T\,. \tag{4.11}$$

Likewise, the second correlation function of the right-hand side of Eq. (4.10) involves only the transverse part of the mass-current-density field. This may then, in principle, be calculated by summing over all possible vortex configurations.

Note that from Eqs. (4.10) and (4.11) it follows that the ρ used to specify H_S in Eq. (4.2) can be interpreted as the macroscopic superfluid density in the absence of vortices.

To sum up the reasoning: The aim is a description of how the thermal fluctuations in the superfluid mass flow destroy the superfluidity. The relevant energy for determining these thermal fluctuations is taken to be the kinetic energy of the superfluid mass flow. It is assumed that this energy, on some suitable length scale, can be described by a H_S of the form of Eq. (4.2). The nontrivial part of this description turns out to be the transverse mass flow, which can be completely specified in terms of vortex configurations.

It turns out that thermally created vortices with vorticity $s=\pm1$ destroy the superfluidity (Berezinskii, 1971; Kosterlitz and Thouless, 1972). Vortices with higher vorticity ($|s|>1$) involve a larger thermal energy as compared to the $s=\pm1$ vortices (see Sec. IV.D.1). Hence $|s|>1$ vortex fluctuations are much rarer events (Kosterlitz, 1974). Furthermore, neutral pairs of vortices with higher vorticity are much harder bound and are hence not directly responsible for the destruction of superfluidity (Kosterlitz and Thouless, 1973). Because of this we shall further simplify the phenomenological description in the following by assuming that only $s=\pm1$ vortices have to be included.

The description of thermal fluctuations for the superfluid in terms of vortices with vorticity $s=\pm1$ is isomophic to the two-dimensional Coulomb gas (Kosterlitz and Thouless, 1973). This connection is reviewed in the following section.

2. Coulomb gas analogy

The analogy between fluctuations in the transverse mass flow and the Coulomb gas model is based on the following observations: Equation (4.7) for the scalar field W may be recognized as the Poisson equation for a positive charge with the charge distribution f_{r_0}. Consequently the field $W(\mathbf{r})$ for a single vortex can be expressed in terms of the Coulomb gas potential $V(\mathbf{r})$ [defined by Eq. (2.1)] as

$$W(\mathbf{r})\Big/\left[s\rho\frac{\hbar}{m^*}\right]=\int d^2r' f_{r_0}(\mathbf{r}-\mathbf{r}')V(\mathbf{r}')\,. \tag{4.12}$$

An arbitrary transverse mass flow is obtained by adding the contributions from all vortices, i.e.,

$$\mathbf{g}_{\perp}(\mathbf{r})=\nabla\times\hat{\mathbf{x}}_3\sum_i s_i W(\mathbf{r}-\mathbf{r}_i)\,, \tag{4.13}$$

where $\mathbf{r}_i$ are the vortex centers, s_i are the vorticities, and the sum i is over all vortices building up the transverse mass flow. It then also follows that $\mathbf{g}_{\perp}(\mathbf{r})$ can be expressed as

$$\mathbf{g}_\perp(\mathbf{r})=\rho\frac{\hbar}{m^*}\nabla\times\hat{\mathbf{x}}_3\int d^2r'\Delta n(\mathbf{r}')V(\mathbf{r}-\mathbf{r}') \tag{4.14}$$

in terms of the Coulomb gas potential $V(\mathbf{r})$ and a Coulomb gas charge density

$$\Delta n(\mathbf{r})=\sum_i s_i f_{r_0}(\mathbf{r}-\mathbf{r}_i)\ , \tag{4.15}$$

provided s_i is interpreted as the charge of a Coulomb gas particle and f_{r_0} as the single-particle charge distribution. The corresponding configuration energy $H_S^{\rm trans}$,

$$H_S^{\rm trans}=\frac{1}{2\rho}\int d^2r\,\mathbf{g}_\perp^2(\mathbf{r})\ , \tag{4.16}$$

can also be expressed in the same quantities by using the Fourier transforms and the convolution theorem

$$H_S^{\rm trans}=\frac{1}{2\rho}\left[\rho\frac{\hbar}{m^*}\right]^2\lim_{k\to 0}(i\mathbf{k}\times\hat{\mathbf{x}}_2)\cdot(-i\mathbf{k}\times\hat{\mathbf{x}}_3)\hat{V}(\mathbf{k})\hat{V}(-\mathbf{k})\Delta\hat{n}(\mathbf{k})\Delta\hat{n}(-\mathbf{k})\ , \tag{4.17}$$

which in real space becomes

$$H_S^{\rm trans}\Big/\left[2\pi\rho\left[\frac{\hbar}{m^*}\right]^2\right]=\tfrac{1}{2}\int d^2r\,d^2r'\Delta n(\mathbf{r})V(\mathbf{r}-\mathbf{r}')\Delta n(\mathbf{r}')\ . \tag{4.18}$$

The right-hand side of Eq. (4.18) can be recognized as the electrostatic energy for a Coulomb gas charge density $\Delta n(\mathbf{r})$. Equations (4.15) and (4.18) constitute the essence of the Coulomb gas analogy.

The point of this analogy is that the results for the two-dimensional Coulomb gas (described in Secs. II and III) can now be mapped over to predictions for the two-dimensional superfluid. Table II gives the mapping between the Coulomb gas model defined in Sec. II.A and the model of a superfluid (defined in Sec. IV.A.1). The main features of the mapping are the following.

The statistical weight of a vortex configuration corresponds to a Coulomb gas configuration. This means that

$$H_{\rm CG}/T^{\rm CG}=H_S^{\rm Trans}/T\ . \tag{4.19}$$

Equations (4.18) and (4.19) give the relation between the "real" temperature T for the superfluid and the Coulomb gas temperature $T^{\rm CG}$, i.e.,

$$T^{\rm CG}=T\Big/\left[2\pi\rho\left[\frac{\hbar}{m^*}\right]^2\right]\ . \tag{4.20}$$

Vorticity corresponds to charge. The Coulomb gas model has $s=\pm 1$. Vortices with vorticity $|s|>1$ are not taken into account. One argument for neglecting $|s|>1$ vortices is that the pair binding strength is proportional to s^2 [compare Eqs. (4.15) and (4.18)].

The single-particle charge density f_{r_0} is determined by the variation of the magnitude of the order parameter ψ close to the center of a vortex with vorticity $|s|=1$ (see Sec. IV.D.1). This vortex core region corresponds to the extension of a single charge.

The creation energy for a vortex corresponds to the electrostatic self-energy [see Eqs. (2.4), (2.6), and (4.18)]. This self-energy takes into account only the contributions associated with $\mathbf{g}_\perp\sim\nabla\theta$. To these must be added the energy associated with the variation of the magnitude of the order parameter ψ in the vortex core region. This energy is proportional to ρ (see Sec. IV.D.1) and, as follows from Eqs. (4.19) and (4.20), it maps over to the additional constant energy E_c in the Coulomb gas model [see Eq. (2.6)].

The leading term of the electrostatic self-energy for a Coulomb gas charge in the limit of large λ_c is $\frac{1}{2}U(0)\simeq\frac{1}{2}\ln(\lambda_c/r_0)$ [see Eq. (2.5)]. The corresponding quantity for the superfluid is $\frac{1}{2}W(0)\simeq\frac{1}{2}\ln(R/r_0)$ [compare Eq. (4.7)], where R is the linear extension of the superfluid. Consequently, non-neutral configurations are possible only for a superfluid with *finite* extension (see Sec. II.A.2). Vortex fluctuations for a superfluid with finite extension are approximately described by the non-neutral Coulomb gas model (defined in Sec. II.A) provided λ_c is identified with the linear extension R (Minnhagen, 1981a).

Vortices may also be induced by rotating the superfluid. If the superfluid is rotated by a constant angular velocity $\omega^{\rm rot}\hat{\mathbf{x}}_3$, then the induced density difference between vortices with vorticity $s=1$ and $s=-1$ is given (see, for example, Wilks, 1967) by

$$\langle\Delta n\rangle=\frac{\omega^{\rm rot}m^*}{\pi\hbar}\ , \tag{4.21}$$

where $\langle\Delta n\rangle$ is the thermal average of the Coulomb gas charge density [see Eq. (4.15)]. As discussed in connection with the non-neutral Coulomb gas (Sec. II.B.2) a nonzero $\langle\Delta n\rangle$ is possible in the large-λ_c limit only if $\Delta E\simeq 2\pi\lambda_c^2\langle\Delta n\rangle$ [see Eq. (2.46), where $2\Delta E$ is the difference in chemical potential between a negative and a positive charge; see also Eq. (2.7)]. The ΔE corresponding to a superfluid follows from Eqs. (2.46), (4.21), and the identification $\lambda_c\simeq R$, i.e.,

$$\Delta E\simeq\frac{R^2\omega^{\rm rot}m^*}{\hbar}\ . \tag{4.22}$$

It corresponds to the difference in creation energy between a vortex with vorticity $s=1$ and $s=-1$. By mul-

TABLE II. Mapping between the two-dimensional Coulomb gas and the neutral superfluid.

Coulomb gas	Coulomb gas	Superfluid
Symbol		
$s=\pm 1$	Charge of a particle	Vorticity of a vortex
$f_{r_0}(r)$	Charge distribution of a particle	Restriction on mass flow close to vortex center; see Eq. (4.7)
T^{CG}	Coulomb gas temperature	$T^{\mathrm{CG}}=T\Big/\left[2\pi\rho\left[\frac{\hbar}{m^*}\right]^2\right]$, where T is the superfluid temperature
Ω/ζ Ω=2D volume ζ=phase-space division	Number of places available for a particle	Number of places available for a vortex
E_c	Nonelectrostatic part of particle self-energy; see Eq. (2.6)	$2\pi\rho\left[\frac{\hbar}{m^*}\right]^2 E_c$=vortex core energy
λ_c	Screening length associated with the background charge; see Eq. (2.3)	R=linear extension of the superfluid; a finite R may be approximately taken into account by $R\simeq\lambda_c$
ΔE	$2\Delta E$=difference in chemical potential between a negative and positive particle; see Eq. (2.7)	$\Delta E\simeq\frac{R^2 m^* \omega^{\mathrm{rot}}}{\hbar}$, ω^{rot}=angular velocity. ΔE corresponds to the energy difference between $s=1$ and $s=-1$ vortices for a rotating superfluid; see Eqs. (2.46), (4.21), (4.22)
$e^{-H_N/T^{\mathrm{CG}}}$ H_N is given by Eq. (2.8)	Boltzmann factor for a particle configuration	Boltzmann factor for the corresponding vortex configuration
Z given by Eq. (2.9)	Partition function for Coulomb gas	Partition function for superfluid

tiplying Eq. (4.11) by the conversion factor $2\pi\rho(\hbar/m^*)^2$ [see Eq. (4.20)], one finds that this energy is proportional to $N_S\hbar\omega^{\mathrm{rot}}$, where N_S is the number of superfluid atoms. From the point of view of the superfluid, the origin of the energy difference is the energy associated with the rotation of the superfluid. This rotation energy is given by $L\omega^{\mathrm{rot}}$, where L is the total angular momentum (see, for example, Wilks, 1967). The energy difference between a positive and negative vortex is just twice this rotation energy per vortex (Wilks, 1967). This may be expressed as $2L_{\mathrm{vortex}}\omega^{\mathrm{rot}}$, where $L_{\mathrm{vortex}}\simeq\rho(\hbar/m)R^2$ is the angular momentum of a single vortex. By use of the conversion factor $2\pi\rho(\hbar/m^*)^2$ this reduces to

$$\Delta E=\frac{L_{\mathrm{vortex}}\omega^{\mathrm{rot}}}{2\pi\rho(\hbar/m^*)^2}\simeq\frac{R^2\omega^{\mathrm{rot}}m^*}{\hbar}\ .$$

The phase-space division ζ for the creation of Coulomb gas charges carries over to the phase-space division for vortices. It is assumed that ζ is of the order of the vortex core and proportional to r_0^2 as for the Coulomb gas (see Sec. II.A.1). It should be noted that for the superfluid this is an additional assumption, albeit intuitively very plausible.

Hence the complete analogy is expressed by the fact that the partition function describing vortex fluctuations for a two-dimensional superfluid can be mapped onto the one describing the two-dimensional Coulomb gas [defined by Eq. (2.9)].

B. Superconductor

The simple phenomenological description of thermodynamic fluctuations for the two-dimensional neutral superfluid can, with minor modifications, be extended to a charged superfluid (Doniach and Hubermann, 1979; Halperin and Nelson, 1979; Minnhagen, 1981a). Precisely as for the neutral case, the superfluid state is characterized by a complex order parameter $\psi(\mathbf{r})=\sqrt{\rho(\mathbf{r})}e^{i\theta(\mathbf{r})}$.

However, the "boson" responsible for the superfluidity now carries a nonzero charge e^* in addition to the mass m^*. In the case of a superconductor this corresponds to the charge and the mass of a Cooper pair, i.e., $e^*=2e$ and $m^*=2m_e$, where e and m_e are the charge and mass of an electron. The finite charge couples to the electromagnetic field. Consequently the relation between the superfluid velocity v_S and the phase of the order parameter θ is modified to [compare Eq. (4.1)]

$$\mathbf{v}_S(\mathbf{r})=\frac{\hbar}{m^*}\nabla\theta(\mathbf{r})-\frac{e^*\mathbf{A}(\mathbf{r},0)}{m^*c}\ , \tag{4.23}$$

where $\mathbf{A}(\mathbf{r},x_3)$ is the transverse gauge vector potential in the three-dimensional space spanned by $\mathbf{x}=(\mathbf{r},x_3)$, the superfluid sheet is taken to be the $x_3=0$ plane (see Fig. 13), and c is the velocity of light. Note that the vector $\mathbf{A}(\mathbf{r},0)$ is parallel to the superfluid sheet, since both $\mathbf{v}_S$ and $\nabla\theta$ are parallel to this sheet by assumption. The mass-current density in the three-dimensional description is given by

$$\mathbf{g}(\mathbf{r},x_3)=\mathbf{g}(\mathbf{r})\delta(x_3)\ , \tag{4.24}$$

where $\mathbf{g}(\mathbf{r})$ is the two-dimensional (areal) mass-current density. For a charged superfluid it is more natural to use the electric current density $\mathbf{j}(\mathbf{r})$ instead of the mass-current density $\mathbf{g}(\mathbf{r})$. They differ by only a trivial factor

$$\mathbf{j}(\mathbf{r})=\frac{e^*}{m^*}\mathbf{g}(\mathbf{r})\ . \tag{4.25}$$

In Sec. IV.B.1 we shall use $\mathbf{g}$ in order to simplify comparison between the cases of a charged and a neutral superfluid. Vortex configurations on a two-dimensional superconductor [i.e., two-dimensional in the sense of Eq. (4.24); the vector potential is, of course, nonzero in the whole three-dimensional space] have been considered by Pearl (1964, 1965), corrections for finite thickness by Clem (1979), vortex magnetic moment by Fetter (1980), and the precise Coulomb gas connection by Minnhagen and Nylén (1985a).

Since the description of vortex fluctuations for a two-dimensional superconductor is entirely analogous to the case of a neutral superfluid, in the following two subsections (IV.B.1 and IV.B.2) we shall only indicate the necessary modifications.

1. Configuration energy

In case of a charged superfluid, the configuration energy H_S corresponding to Eq. (4.2) is given by

$$H_S=\int d^2r\left[\frac{\mathbf{g}^2(\mathbf{r})}{2\rho}+\frac{e^*\mathbf{A}(\mathbf{r},0)\cdot\mathbf{g}(\mathbf{r})}{2m^*c}\right]\ . \tag{4.26}$$

The first term on the right-hand side is the kinetic energy of the mass flow, precisely as for the neutral case. The second term is a representation of the induced magnetic field energy

$$\int d^2r\,dx_3\frac{\mathbf{B}^2(\mathbf{r},x_3)}{8\pi}\ .$$

By a partial integration and by use of the Maxwell equation $\mathrm{curl}(\mathrm{curl}\,\mathbf{A})=(4\pi/c)\mathbf{j}$ together with Eq. (4.25), we can express this magnetic field energy in terms of the mass flow $\mathbf{g}$ on the superfluid sheet, as in Eq. (4.26).

In parallel with the neutral case (Sec. IV.A.1), the mass current is conveniently split into longitudinal and transverse parts, which are expressible in terms of the scalar fields $\Phi(\mathbf{r})$ and $W(\mathbf{r})$, respectively. Equation (4.4a) is unchanged, while Eq. (4.4b) is modified to

$$\mathbf{g}_\perp(\mathbf{r})=\nabla\times\hat{\mathbf{x}}_3W(\mathbf{r})-\frac{\rho e^*}{m^*c}\mathbf{A}(\mathbf{r},0) \tag{4.27}$$

because the vector potential $\mathbf{A}(\mathbf{r},0)$ is a transverse vector field in the plane of the superfluid. The scalar field $W(\mathbf{r})$ describes the part of the transverse current associated with the gradient of the phase of the order parameter [compare Eq. (4.23)]. The definition of a vortex is unaltered, and the single-valuedness of the order parameter again leads to the restriction on W expressed by Eq. (4.7).

Just as for the neutral case, H_S separates into longitudinal and transverse parts given by Eq. (4.8). In the same way, the definition of the superfluid mass density ρ_S, given by Eq. (4.9), again leads to the definition in terms of correlation functions given by Eq. (4.10).

There is a complete analogy between a finite rotation, $\omega^{\mathrm{rot}}\hat{\mathbf{x}}_3$, for a neutral superfluid (see Sec. IV.A.2) and an external magnetic field, $B_{\mathrm{ex}}\hat{\mathbf{x}}_3$, for a charged superfluid (see, for example, Vinen, 1969). The energy associated with the magnetic field B_{ex} is given by

$$H_S^{\mathrm{ex}}=-MB_{\mathrm{ex}}\ , \tag{4.28}$$

where M is the total magnetic moment of the superfluid. The total magnetic moment is given by

$$\mathbf{M}=\frac{1}{2c}\int d^2r[r\times\mathbf{j}(\mathbf{r})]$$
$$=\frac{e^*}{2cm^*}\int d^2r[\mathbf{r}\times\mathbf{g}_\perp(\mathbf{r})]=\frac{e^*}{2cm^*}\mathbf{L}\ , \tag{4.29}$$

where $\mathbf{L}$ is the total angular momentum of the superfluid. The second equality in Eq. (4.29) follows from Eq. (4.25) and from the condition $\Phi=\mathrm{const}$ on the boundary of the superfluid sheet [compare Eq. (4.4a)]. Comparing Eqs. (4.28) and (4.29) with the corresponding rotational energy $L\omega^{\mathrm{rot}}$, in the case of the neutral superfluid (see Sec. IV.A.2), leads to the correspondence

$$\omega^{\mathrm{rot}}=\omega^{\mathrm{cycl}}/2\ , \tag{4.30a}$$

where ω^{cycl} is the cyclotron frequency

$$\omega^{\mathrm{cycl}}=\frac{e^*B_{\mathrm{ex}}}{m^*c}\ . \tag{4.30b}$$

This analogy means that, within a Coulomb gas description of vortices, a perpendicular magnetic field for a charged superfluid is equivalent to a rotation velocity $\omega^{\mathrm{rot}}=\omega^{\mathrm{cycl}}/2$ for a corresponding neutral superfluid.

2. Coulomb gas analogy

Just as for the neutral superfluid, the Coulomb gas analogy for the charged superfluid is based on the Poisson equation [Eq. (4.7)] for the scalar field $W(\mathbf{r})$. This equation corresponds to a single vortex. An arbitrary transverse mass flow is specified by a configuration of vortices. In analogy with Eq. (4.13), it is given by

$$\mathbf{g}_\perp(\mathbf{r})=\nabla\times\hat{\mathbf{x}}_3\sum_i s_i W(\mathbf{r}-\mathbf{r}_i)-\frac{\rho e^*}{m^*c}\mathbf{A}(\mathbf{r},0)\ , \quad (4.31a)$$

where the vector potential $\mathbf{A}$ is determined from the Maxwell equation

$$\nabla\times[\nabla\times\mathbf{A}(\mathbf{r},0)]=\frac{e^*4\pi}{m^*c}\mathbf{g}_\perp(\mathbf{r})\ . \quad (4.31b)$$

From the Fourier transforms of Eqs. (4.7), (4.31a), and (4.31b) one obtains

$$\hat{\mathbf{g}}_\perp(\mathbf{k})=2\pi\frac{\rho\hbar}{m^*}\frac{\Delta\hat{n}(k)}{k+u(k)\Lambda^{-1}}\ , \quad (4.32a)$$

where

$$\Lambda=\frac{1}{2\pi}\left[\frac{m^*c}{e^*}\right]^2\frac{1}{\rho}\ , \quad (4.32b)$$

$u(k)$ is a dimensionless function of k, and $u(0)=1$ (Minnhagen and Nylén, 1985a). $\Delta\hat{n}$ is the Fourier transform of the Coulomb gas charge density [see Eq. (4.15)]. In the special case $f_{r_0}(\mathbf{r})=\delta(\mathbf{r})$ [compare Eqs. (4.6) and (4.7)], the function $u(k)$ is identically equal to one (Pearl, 1965). The length Λ, given by Eq. (4.32b), may be interpreted as a magnetic penetration depth (Pearl, 1965), i.e., it gives the scale of penetration of a magnetic field perpendicular to the superconducting sheet. It is related to the London penetration depth Λ_L for a three-dimensional superconductor by $\Lambda=2\Lambda_L^2/d$, where d is the thickness of the "two-dimensional" superconductor (Pearl, 1965; see Fig. 13). The transverse part of the configuration energy H_S [compare Eq. (4.26)] is given by

$$H_S^{\rm trans}=\int d^2r\left[\frac{\mathbf{g}_\perp^2(\mathbf{r})}{2\rho}+\frac{e^*\mathbf{A}(\mathbf{r},0)\cdot\mathbf{g}_\perp(\mathbf{r})}{2m^*c}\right]. \quad (4.33)$$

In analogy with Eq. (4.18), H_S given by Eq. (4.33) can be expressed as (Minnhagen and Nylén, 1985a)

$$H_S^{\rm trans}\Big/\left[2\pi\rho\left[\frac{\hbar}{m^*}\right]^2\right]=\tfrac{1}{2}\int d^2r\,d^2r'\Delta n(\mathbf{r})Q(\mathbf{r}-\mathbf{r}')\Delta n(\mathbf{r}')\ , \quad (4.34)$$

where

$$\hat{Q}(k)=\frac{2\pi}{k}\frac{k+\Lambda^{-1}}{[k+u(k)\Lambda^{-1}]^2}\ . \quad (4.35)$$

In the limit $\Lambda=\infty$, the Fourier transform of the potential Q given by Eq. (4.35) reduces to the Fourier transform of the Coulomb gas potential $\hat{V}(k)=2\pi/k^2$ and, consequently, Eq. (4.34) reduces to the corresponding equation for the neutral superfluid [Eq. (4.18)]. In other words, in the limit of a large magnetic penetration depth Λ, the induced magnetic energy,

$$\int d^2r\,dx_3\frac{B^2(\mathbf{r},x_3)}{8\pi}\ ,$$

can be neglected in comparison with the kinetic mass flow energy, $\int d^2r\,g^2(\mathbf{r})/2\rho$ [see Eqs. (4.26) and (4.33); the second term on the right-hand side of these equations represents the induced magnetic field energy, as explained below Eq. (4.26)]. Consequently Eq. (4.18) constitutes the essence of the Coulomb gas analogy also for a charged superfluid in the limit of large Λ. We shall in the following restrict the discussion of charged superfluids to this "large-Λ case" when the induced magnetic field energy can be neglected.

As first pointed out by Beasley *et al.* (1979), thin, "dirty" type-II superconducting films may typically have magnetic penetration depths of the order of centimeters and hence of the order of a typical sample size. For this type of physical realization of a charged superfluid, the Coulomb gas analogy of vortex fluctuations should apply (Beasley *et al.*, 1979). Comparison between experiments on such superconducting films and the Coulomb gas prediction are discussed in Sec. V.B.

Since the limit of large Λ essentially reduces the charged superfluid to the neutral one, the Coulomb gas analogy given in Table II remains unaltered with two minor modifications. These modifications are given in Table III. The first concerns the identification of λ_c: the electrostatic self-energy for a Coulomb gas is given by $\frac{1}{2}U(0)\simeq\frac{1}{2}\ln(\lambda_c/r_0)$ [see Eq. (2.5)]. The corresponding energy for a charged superfluid is given by (Pearl, 1965; Minnhagen, 1981a)

$$\tfrac{1}{2}Q(0)\simeq\begin{cases}\frac{1}{2}\ln(R/r_0), & R<\Lambda\\ \frac{1}{2}\ln(\Lambda/r_0), & R>\Lambda\ ,\end{cases} \quad (4.36)$$

where R is the linear extension of the superconductor. Thus, in analogy with the neutral case (see Sec. IV.A.2), the Coulomb gas model approximately describes vortex fluctuations for the charged "large-Λ" superfluid provided λ_c is identified with the smaller of Λ and R. The second modification concerns an external perpendicular magnetic field $B_{\rm ex}$. By analogy with a finite rotation of a neutral superfluid [see Eqs. (4.30)] one obtains [with the aid of Eqs. (2.46), (4.21), (4.32b), and the conversion factor $2\pi\rho(\hbar/m^*)^2$] the energy difference between a positive and negative vortex,

$$2\pi\rho\left[\frac{\hbar}{m^*}\right]^2\Delta E=m_v B_{\rm ex}\ , \quad (4.37a)$$

where

$$m_v=\frac{\lambda_c^2\varphi_0}{2\pi\Lambda} \quad (4.37b)$$

TABLE III. Mapping between the two-dimensional Coulomb gas and a "large-Λ" superconductor. The mapping is given by Table II with the listed modifications.

Coulomb gas	Coulomb gas	Superconductor
Symbol		
λ_c	Screening length associated with background charge; see Eq. (2.3)	$\lambda_c \sim \begin{cases} R & \text{if } \Lambda > R \\ \Lambda & \text{if } \Lambda < R \end{cases}$ R=linear extension Λ=magnetic penetration depth
ΔE	$2\Delta E$=difference in chemical potential between negative and positive particle; see Eq. (2.7)	$\Delta E = \frac{2\pi\lambda_c^2 B_{ex}}{\varphi_0}$, where B_{ex} is an external perpendicular magnetic field; see Eq. (4.37a)

and where $\varphi_0 = hc/2|e|$ is the flux quantum. The quantity m_v may be interpreted as the magnetic moment for a single positive vortex within the Coulomb gas description (Minnhagen, 1981a).

For example, consider the case when $\Lambda \gg R$. As discussed above, the Coulomb gas model in this case approximately describes vortex fluctuations for a "large-Λ" superconducting film provided λ_c is chosen to be of the order of R. From Eq. (4.37b) one then obtains $m_v \simeq R^2\varphi_0/2\pi\Lambda$, which correctly reproduces the result for the magnetic moment of a single vortex on a superconducting film (Fetter, 1980).

The Coulomb gas charge density corresponding to a perpendicular external magnetic field can be obtained from Eq. (4.21) and the rotation analogy given by Eqs. (4.30),

$$\langle \Delta n \rangle = B_{ex}/\varphi_0 \ . \tag{4.38}$$

Equation (4.38) expresses the fact that the flux per unit area, B_{ex}, passing through a superconducting film is equal to the average vorticity per unit area times the flux quantum.

C. Superfluid density and the universal jump

The fundamental characteristic of the transition between the superfluid and normal state is the vanishing of the macroscopic superfluid density. The superfluid density can, as discussed in connection with Eq. (4.10), be expressed in terms of mass-current correlation functions. The longitudinal part of this expression can be directly integrated out [see Eq. (4.11)], leaving only the transverse part. The transverse part is completely specified by vortex configurations. The vortex-Coulomb gas analogy then allows one to express the superfluid density as a Coulomb gas correlation function. By translational invariance one has

$$\int d^2r \langle \mathbf{g}_\perp(\mathbf{r})\cdot\mathbf{g}_\perp(0)\rangle = \lim_{\mathbf{k}\to 0} \langle \hat{\mathbf{g}}_\perp(\mathbf{k})\cdot\hat{\mathbf{g}}_\perp(\mathbf{k})\rangle \ , \tag{4.39}$$

which via the convolution theorem applied to Eq. (4.14) becomes

$$\left[\rho\frac{\hbar}{m^*}\right]^2 \lim_{\mathbf{k}\to 0} (i\mathbf{k}\times\hat{\mathbf{x}}_3)\cdot(-i\mathbf{k}\times\hat{\mathbf{x}}_3)\hat{V}(\mathbf{k})\hat{V}(-\mathbf{k}) \times \langle \Delta\hat{n}(\mathbf{k})\Delta\hat{n}(-\mathbf{k})\rangle \ . \tag{4.40}$$

Since $\hat{V}(\mathbf{k}) = 2\pi/k^2$ [compare Eq. (2.1)], the result for the transverse part of Eq. (4.10) is

$$\int d^2r \langle \mathbf{g}_\perp(\mathbf{r})\cdot\mathbf{g}_\perp(0)\rangle = 2\pi \left[\rho\frac{\hbar}{m^*}\right]^2 \times \lim_{\mathbf{k}\to 0} \hat{V}(\mathbf{k})\langle \Delta\hat{n}(\mathbf{k})\Delta\hat{n}(-\mathbf{k})\rangle \ . \tag{4.41}$$

Inserting this result, together with Eqs. (4.11) and (4.20), into the definition of the superfluid density, Eq. (4.10), then gives

$$\frac{\rho_S}{\rho} = \lim_{\mathbf{k}\to 0}\left[1 - \frac{\hat{V}(k)}{T^{CG}}\langle \Delta\hat{n}(\mathbf{k})\Delta\hat{n}(-\mathbf{k})\rangle\right] . \tag{4.42}$$

The left-hand side may be recognized as the inverse of the dielectric constant $\varepsilon_\infty = \hat{\varepsilon}(k=0)$, in which $\hat{\varepsilon}(k=0)$ is the Fourier transform of the dielectric function for the Coulomb gas at $\mathbf{k}=0$ [compare Eqs. (3.3c) and (3.4)]. Thus there exists a simple direct relation between the superfluid density and the dielectric constant of the Coulomb gas (Myerson, 1978; Minnhagen and Warren, 1981),

$$\frac{\rho_S}{\rho} = \frac{1}{\varepsilon_\infty} \ . \tag{4.43}$$

The Coulomb gas quantity $1/(T^{CG}\varepsilon_\infty)$ jumps from a finite value to zero at the Kosterlitz-Thouless transition [compare Eq. (2.36) and the paragraph following this equation]. According to the Kosterlitz renormalization-group equations, the size of the jump is universal (that is, temperature and fugacity independent) and equal to 4. As discussed in Sec. III.D, the result may be expected to hold provided the fugacity z at the transition is smaller than some critical value (see Fig. 12). By aid of Eqs. (4.43) and

(4.20) one easily translates this universal Coulomb gas jump into superfluid variables,

$$\frac{\rho_S}{T}=\begin{cases}\dfrac{2k_B m^{*2}}{\pi\hbar^2} & \text{for } T\to T_c^- ,\\ 0 & \text{for } T\to T_c^+ ,\end{cases} \tag{4.44}$$

where the Boltzmann constant k_B has been reinserted for completeness. This is the celebrated universal jump prediction for the superfluid density of a two-dimensional superfluid (Nelson and Kosterlitz, 1977). The important feature is that the size of the jump of the quantity ρ_S/T involves only fundamental constants and the mass of the "boson." For an ideal ^{4}He film this is just the mass of the ^{4}He atom, while for a superconductor it is twice the electron mass. The prediction has been verified by experiments, as will be discussed in Sec. V.A.

From the Coulomb gas point of view, the fact that the universal jump is borne out by experiments on superfluid films can be taken as an *ipso facto* verification that the Coulomb gas fugacities corresponding to these experimental realizations are small enough for the universal jump prediction to be valid.

We shall now illustrate the Coulomb gas—superfluid film analogy further by considering how the superfluid density is modified when a neutral superfluid film is rotated or a perpendicular magnetic field is applied to a superconducting film. As discussed in Secs. IV.A.2 and IV.B.2, we can model this situation approximately by the non-neutral Coulomb gas with an appropriate choice of λ_c (compare Tables II and III). Defining the superfluid density by Eq. (4.9b) leads, in the case of the non-neutral Coulomb gas, to (Minnhagen, 1981a).

$$\frac{\rho_S}{\rho}=\lim_{k\to k_{\min}}\left[\frac{1}{\tilde{\varepsilon}}\left[\frac{k^2+\lambda_c^{-2}}{k^2+\lambda^{-2}}\right]\right] , \tag{4.45}$$

where $k_{\min}\sim 1/R$ is introduced instead of $k=0$ because R, the size of the sample, is the macroscopic length scale associated with ρ_S. Below T_c the screening length λ is directly related to the charge imbalance $\langle \Delta n \rangle$, i.e., $\lambda^{-2}=2\pi\langle \Delta n \rangle/\tilde{\varepsilon}T^{\mathrm{CG}}$ [compare Eq. (2.21b) and Sec. II.B.2]. Thus Eq. (4.45), together with Eqs. (4.20) and (4.21), suggests for the neutral rotating superfluid below T_c

$$\frac{\rho_S}{\rho}=\frac{1}{\varepsilon_\infty}\frac{1}{1+\omega^{\mathrm{rot}}/\omega_0^{\mathrm{rot}}} , \tag{4.46a}$$

where

$$\omega_0^{\mathrm{rot}}\simeq\frac{k_B T m^*}{h\rho R^2} . \tag{4.46b}$$

Note that for $\omega^{\mathrm{rot}}=0$, Eq. (4.46a) reduces to Eq. (4.43). For $\omega^{\mathrm{rot}}\neq 0$ the Coulomb gas analogy predicts that the superfluid density decreases with rotation velocity ω^{rot} and that, for large velocities compared to a characteristic velocity ω_0^{rot}, the superfluid density is proportional to $(\omega^{\mathrm{rot}})^{-1}$. Equation (4.46b) gives an order-of-magnitude estimate of the characteristic rotation velocity ω_0^{rot}.

The corresponding $T<T_c$ result for the "large-Λ" superconductor in an external magnetic field is obtained by using the analogy given by Eqs. (4.30):

$$\frac{\rho_S}{\rho}=\frac{1}{\varepsilon_\infty}\frac{1}{1+B_{\mathrm{ex}}/B_0} , \tag{4.47a}$$

where

$$B_0\simeq\frac{k_B T\Lambda}{\varphi_0[\min(R,\Lambda)]^2} . \tag{4.47b}$$

The interpretation is completely analogous to that of Eqs. (4.46).

More generally, the result from the Coulomb gas—superfluid analogy for the superfluid density may be summarized by

$$\frac{\rho_S}{\rho}=\frac{1}{\tilde{\varepsilon}}\frac{1}{1+(\lambda_{\mathrm{exp}}/\lambda_F)^2} , \tag{4.48}$$

where λ_{exp} is a length of the order of the length scale probed by a specific experiment for the superfluid, λ_F is the Coulomb gas screening length associated with free Coulomb gas charges, i.e., $\lambda_F^{-2}=2\pi n_F/\tilde{\varepsilon}T^{\mathrm{CG}}$ [compare Eq. (2.21b)], and $\tilde{\varepsilon}$ is a dielectric constant describing the polarization due to bound pairs [defined by Eq. (2.21a)]. The density of free charges is given by Eq. (2.44), and Coulomb gas variables are translated into superfluid variables with the aid of Tables II and III.

D. Ginzburg-Landau Coulomb gas

In Secs. IV.A and IV.B we have reviewed how a simple phenomenological description of vortex fluctuations for superfluid films leads to a model that is isomorphic to the two-dimensional Coulomb gas. However, this phenomenological description of superfluid films is still incomplete, in as far as some quantities have been left unspecified. These quantities are the "bare" superfluid density ρ [see Eq. (4.2)], the vortex core function $f_{r_0}(r)$ [see Eq. (4.7)], the vortex core energy E_c (see Table II), and the phase-space division ζ (see Table II). Section IV.D.1 gives a brief outline of how ρ, f_{r_0}, and E_c can be obtained within a Ginzburg-Landau approximation. The resulting phenomenological description of vortex fluctuations for superfluid films is referred to as the Ginzburg-Landau Coulomb gas model. This model leads to certain scaling relations, referred to as Coulomb gas scaling (Minnhagen, 1981b). Coulomb gas scaling is described in Sec. IV.D.2.

1. Definition

In this section the configuration energy H_S is phenomenologically connected to a Ginzburg-Landau-type approximation. The starting point is the Ginzburg-Landau form of the free energy, ΔG_S, associated with the order parameter ψ,

$$\Delta G_S = \int d\mathbf{x} \left[\alpha(T) |\psi|^2 + \frac{\beta}{2} |\psi|^4 + \frac{B^2}{8\pi} - \frac{\mathbf{B}\cdot\mathbf{B}_{\text{ext}}}{4\pi} + \frac{1}{2(m^*)^2} \left| \left[-i\hbar \frac{d}{d\mathbf{x}} + \frac{e^*}{c}\mathbf{A} \right] \psi \right|^2 \right], \tag{4.49}$$

where β is a phenomenological constant greater than zero, $\alpha(T)$ is a phenomenological function of T that is positive (negative) for $T > T_{c0}$ ($T < T_{c0}$). T_{c0} is the phenomenological Ginzburg-Landau temperature. The integral is over all space, $\mathbf{x}=(\mathbf{r},x_3)$, and $d/d\mathbf{x}$ denotes the gradient. For a review of the Ginzburg-Landau theory see, for example, Lifshitz and Pitaevski (1980). The Ginzburg-Landau approximation consists of estimating the free energy with the minimum of ΔG_S with respect to variations of ψ and $\mathbf{A}$, subject to appropriate boundary conditions.

In the absence of currents and magnetic fields the minimum of ΔG_S corresponds to a position-independent order parameter ψ given by

$$\rho_0(T) = |\psi|^2 = \begin{cases} \dfrac{|\alpha(T)|}{\beta} & \text{for } T < T_{c0}, \\ 0 & \text{for } T > T_{c0}, \end{cases} \tag{4.50}$$

where ρ_0 is the superfluid mass density. Another important quantity within this description is the Ginzburg-Landau coherence length ξ,

$$\xi = \frac{\hbar}{m^*} \frac{1}{\sqrt{2|\alpha(T)|}} . \tag{4.51}$$

This length gives the scale of variation of ψ (see, for example, Lifshitz and Pitaevskii, 1980). Consequently, from the Ginzburg-Landau point of view, a superfluid becomes effectively two dimensional when the thickness of the superfluid sheet is everywhere much smaller than the coherence length ξ.

The object is to use a Ginzburg-Landau-type approximation for the energies of vortex configurations. Obviously, the zero-vortex configuration corresponds, within a Ginzburg-Landau approximation, to the constant order parameter ρ_0 given by Eq. (4.50). Furthermore this Ginzburg-Landau ρ_0 can be identified with the "bare" superfluid density ρ, used to specify H_S [see Eqs. (4.2) and (4.26)]. This is because ρ in H_S can be interpreted as precisely the superfluid density in the absence of vortices [see the paragraph just below Eq. (4.11)].

The energy of a one-vortex configuration is obtained by minimizing ΔG_S given by Eq. (4.49), subject to the constraint of having precisely one vortex. Far from the vortex center the superfluid density corresponding to a single vortex, $\rho_1(r)$, approaches the constant zero-vortex value ρ_0. The ratio $\rho_1(r)/\rho_0$ is plotted in Fig. 14. Note that the region where $\rho_1(r)$ deviates appreciably from ρ_0 is given by the Ginzburg-Landau coherence length ξ. In other words, ξ is a measure of the linear extension of the vortex core region.

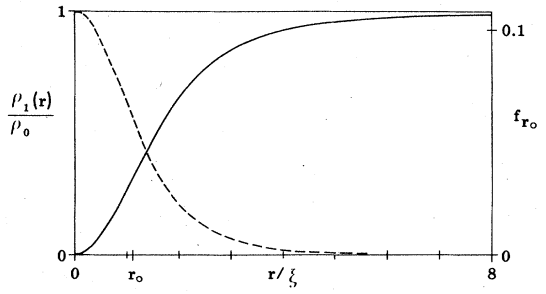

FIG. 14. Variation of the superfluid density close to the vortex center for a vortex with vorticity $|s|=1$ and the single-particle charge distribution of the corresponding Coulomb gas particle. The solid curve is the ratio between the superfluid density for a one-vortex configuration $\rho_1(r)$ and the zero-vortex configuration ρ_0, plotted as a function of the distance r from the vortex center (in units of the Ginzburg-Landau coherence length ξ). The dashed curve is the corresponding single-particle charge distribution $f_{r_0}(r)$. The length r_0 is a measure of the linear extension of the charge distribution (r_0 is defined in connection with Fig. 15).

The Ginzburg-Landau energy corresponding to a single vortex has two distinct contributions (Pearl, 1965). One part is associated with the gradient of the phase of the order parameter, $\nabla\theta$. This is the part accounted for by H_S [see Eqs. (4.2) and (4.26)]. The second part is associated with the variation of the magnitude of the order parameter close to the vortex center. For a single vortex at the origin with vorticity s, this additional part is given by (Minnhagen and Nylén, 1985a)

$$2\pi \left[\frac{\hbar}{m^*} \right]^2 \rho_0 \int dr \frac{r}{4} \left\{ \left[1 - \frac{\rho_1(r)}{\rho_0} \right]^2 \Big/ \xi^2 + \frac{\rho_0}{2\rho_1(r)} \left[\frac{d}{dr} \left[\frac{\rho_1(r)}{\rho_0} \right] \right]^2 \right\} . \tag{4.52}$$

Comparison with Table II identifies the dimensionless integral of Eq. (4.52) as E_c. For a vortex with $|s|=1$, the constant E_c is in the limit $\Lambda/\xi \to \infty$ equal to 0.390 (Hu, 1972). In comparison, the corresponding value for a vortex with $|s|=2$ is 1.208 (Minnhagen and Nylén, 1985a). In the Coulomb gas model, $2E_c$ corresponds to the "threshold energy" of a neutral pair [that is, the smallest energy needed to create a neutral pair; compare Eq. (2.14)]. Consequently neutral pairs of vortices with vorticity $|s|=2$ (and more generally vortices with vorticity $|s|>1$) are very rare compared to pairs of vortices with vorticity $|s|=1$.

The charge distribution $f_{r_0}(r)$ of a Coulomb gas particle corresponding to a vortex with vorticity $|s|=1$ (see Table II) is related to the superfluid density for a single vortex, ρ_1, by (Minnhagen and Nylén, 1985a)

$$f_{r_0}(r)=\frac{1}{2\pi r\rho_0}\frac{d}{dr}\rho_1(r)\ . \tag{4.53}$$

The single-particle charge distribution f_{r_0} is plotted in Fig. 14 together with the superfluid density for a single vortex $\rho_1(r)$, to which it is related via Eq. (4.53). Also shown in the figure is the quantity r_0, which is a measure of the linear extension of a Coulomb gas charge corresponding to a single vortex. The precise definition of r_0 is explained below in connection with Eq. (4.55).

The energy of one vortex at the origin differs from the free energy associated with one vortex because of the entropy contribution S, due to the freedom of placing the vortex anywhere. This additional contribution to the free energy is given by $-TS$. The entropy S can be expressed as $S=\ln(\Omega/\zeta)$ where Ω is the total area of the superfluid sheet and ζ is the phase-space area associated with one vortex. It is assumed in the phenomenological description that the phase-space area ζ is proportional to the vortex core region ξ^2. Since ξ is proportional to r_0, it follows that ζ can be expressed as $\zeta=\tilde{c}r_0^2$ where $\tilde{c}$ is a temperature-independent constant [compare the definition of r_0 given in connection with Eq. (4.55)]. This definition of ζ is in accordance with the Coulomb gas model (see Sec. II.A.1).

Configurations with more than one vortex can, to good approximation, be regarded as superpositions of one-vortex configurations, provided the vortex separation is large enough compared to the vortex core extension ξ (Pearl, 1965). This follows because the current contributions from vortices are additive [see Eq. (4.13)]. Consequently, if we use a phenomenological Ginzburg-Landau description as the starting point, then the vortex configurations are described by the Coulomb gas model with specific Ginzburg-Landau estimates of the Coulomb gas parameters. This is summarized in Table IV.

The interaction between the vortices in a neutral pair is, within the Ginzburg-Landau Coulomb gas model [compare Eq. (2.4)], given by

$$U(r)-U(0)=-\int d^2r'd^2r''[f_{r_0}(|\mathbf{r}-\mathbf{r}'|)-f_{r_0}(r')]\times V(|\mathbf{r}'-\mathbf{r}''|)f_{r_0}(r'')\ , \tag{4.54}$$

where f_{r_0} is given by Eq. (4.53). $U(r)-U(0)$ is plotted in Fig. 15. For $r>r_0$ the interaction rapidly approaches a logarithmic asymptote given by

$$U(r)-U(0)=\ln\left[\frac{r}{2r_0}\right]\ , \tag{4.55}$$

where $r_0=1.12\xi$ (Minnhagen and Nylén, 1985a).

The phenomenological Ginzburg-Landau Coulomb gas description presumes that a function $\alpha(T)$ can be extracted from the experiment [compare Eqs. (4.49) and (4.50)]. In practice the first term in an expansion around T_{c0} may be sufficient, i.e.,

$$\alpha(T)=\alpha'\frac{T-T_{c0}}{T_{c0}}\ . \tag{4.56}$$

TABLE IV. Ginzburg-Landau Coulomb gas. (For definition of Coulomb gas quantities see Table II.)

Coulomb gas quantity	Ginzburg-Landau quantity
E_c	0.390, See Eq. (4.52)
f_{r_0}	$\frac{1}{2\pi r\rho_0}\frac{d}{dr}\rho_1(r)$, See Eq. (4.53)
r_0	1.12ξ, See Fig. 15
ζ	$\tilde{c}r_0^2$, $\tilde{c}$ is a constant

The standard phenomenological assumption in a Ginzburg-Landau description is that Eq. (4.56) holds close enough to T_{c0} (see, for example, Lifshitz and Pitaevski, 1980).

2. Coulomb gas scaling

The Ginzburg-Landau Coulomb gas model implies the existence of scaling relations for superfluid films (Minnhagen, 1981b). Among the relations between quantities characterizing a superfluid film, those which can be expressed entirely in Coulomb gas quantities should be identical for all superfluid films describable by the Ginzburg-Landau Coulomb gas model. Examples of such scaling relations are given in the present section.

The point of the Coulomb gas scaling relations is that they are rather directly open to experimental verification, as will be shown in Sec. V.

As a first example of a Coulomb gas scaling relation, let us consider a superconducting film in the absence of an external magnetic field. This is described by a

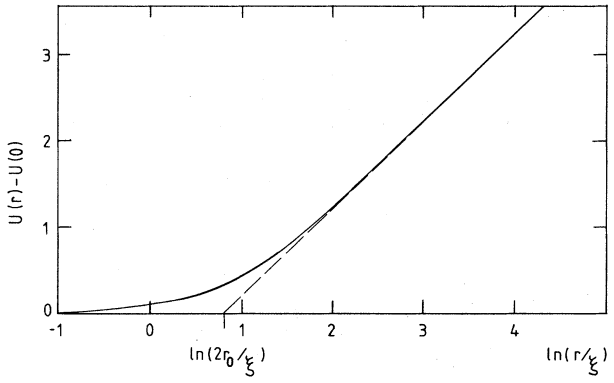

FIG. 15. The interaction energy between two vortices with vorticities $s=1$ and $s=-1$, respectively. The interaction energy $U(r)-U(0)$ is plotted against $\ln(r/\xi)$, i.e., the logarithm of the separation between the vortices in units of the Ginzburg-Landau coherence length ξ (solid curve). For large r the $U(r)-U(0)$ curve rapidly approaches a straight line (dashed line). This line is given by $\ln(r/2r_0)$. The crossing with the horizontal axis gives $r_0\simeq 1.12\xi$. The length r_0 can be regarded as a measure of the linear extension of the single-particle charge distribution.

Ginzburg-Landau Coulomb gas with $\Delta E=0$ and $\lambda_c=\infty$ (compare Tables II and III). A Coulomb gas quantity can then only be a function of the variables (ξ, T^{CG}) (compare Tables II and IV). Consequently a dimensionless Coulomb gas quantity can only be a function of the dimensionless variable T^{CG} [compare Eqs. (4.20) and (4.50)], where

$$T^{CG}=\frac{T}{\alpha(T)}\frac{\beta(m^*)^2}{2\pi\hbar^2} . \tag{4.57}$$

An example of such a dimensionless Coulomb gas quantity is $\xi^2 n_F$, where n_F is the density of free vortices. According to the Bardeen-Stephen formula, this quantity is proportional to the flux-flow resistance (Bardeen and Stephen, 1965), i.e.,

$$\frac{R}{R_N}=2\pi\xi^2 n_F , \tag{4.58}$$

where R is the flux-flow resistance and R_N is the normal-state resistance. The relation given by Eq. (4.58) may alternatively be viewed as an empirical fact (see, for example, Kim and Stephen, 1969). It then follows that the resistance ratio R/R_N as a function of T^{CG} should be a "universal" (i.e., sample-independent) function of the variable T^{CG} for all superconducting films describable by the Ginzburg-Landau Coulomb gas model (Minnhagen, 1981b). A convenient scaling variable is given by

$$X(T)=T^{CG}(T)/T^{CG}(T_c) , \tag{4.59a}$$

which reduces to

$$X(T)=\frac{T}{T_{c0}-T}\frac{T_{c0}-T_c}{T_c}=\frac{\frac{T_{c0}}{T_c}-1}{\frac{T_{c0}}{T}-1} \tag{4.59b}$$

when the approximation given by Eq. (4.56) applies (Minnhagen, 1981b). Experimental evidence for the prediction that the resistance ratio R/R_N is a sample-independent function of X is discussed in Sec. V.B.1.

Next we consider a superconducting film in an external perpendicular magnetic field, B_{ex}. The corresponding Coulomb gas scaling for the resistance can be obtained in the following way (Minnhagen, 1983a): According to the non-neutral Coulomb gas (described in Sec. II.B.2), the dimensionless quantity $\xi^2 n_F$ can, in the limit $\lambda_c\to\infty$, only be a function of the two dimensionless variables T^{CG} and $\xi^2\langle\Delta n\rangle$ [compare Eq. (2.44)]. The difference in density between positive and negative Coulomb gas charges, $\langle\Delta n\rangle$, is related to B_{ex} by Eq. (4.38), i.e., $B_{ex}=\varphi_0\langle\Delta n\rangle$. It follows that the resistance ratio R/R_N should be a "universal" function of the two variables X and B_{ex}/B_{sc}, where X is given by Eqs. (4.59) and the scaling magnetic field B_{sc} is given by (Minnhagen, 1983a)

$$B_{sc}=\varphi_0/(2\pi\xi^2) . \tag{4.60}$$

A constant (two-dimensional) current density j_{ex}, imposed across a superconducting film, gives rise to a Lorentz force F acting on the vortex, i.e.,

$$F=sj_{ex}\varphi_0/c , \tag{4.61}$$

where φ_0 is the flux quantum, c the velocity of light, and s the vorticity (see, for example, Tinkham, 1975). The direction of the Lorentz force is opposite for a $s=1$ vortex from that for a $s=-1$ vortex. In the Coulomb gas analogy the Lorentz force acting on the vortices corresponds to an electric field acting on the Coulomb gas particles (Myerson, 1978; compare Secs. II.B.2). Going through the Coulomb gas analogy [which in this case amounts to comparing Eqs. (4.61), (2.48), (4.19), and (4.20)] gives

$$D=\xi^{-1}j_{ex}/j_{sc} , \tag{4.62a}$$

where

$$j_{sc}=\frac{c}{\varphi_0\xi}2\pi\rho\left[\frac{\hbar}{m^*}\right]^2=\frac{\varphi_0 c}{4\pi^2\Lambda\xi} . \tag{4.62b}$$

Here D is an external electric field applied to the Coulomb gas, and j_{sc} plays the role of a scaling current. In other words, an imposed constant current j_{ex} corresponds to the dimensionless Coulomb gas variable j_{ex}/j_{sc}. This means, for example, that the resistance ratio R/R_N within a Ginzburg-Landau Coulomb gas description is a "universal" function of the variables X, B_{ex}/B_{sc}, and j_{ex}/j_{sc} (Minnhagen, 1984a).

In the case $B_{ex}=0$, the density of free vortices, generated by the imposed current j_{ex} for temperatures below T_c, can be estimated by use of Eqs. (2.51) and (4.62a). In the limit of small j_{ex}/j_{sc} one obtains (Halperin, 1979)

$$R/R_N\sim\xi^2 n_F\sim(D\xi)^{1/2T^{CG}\varepsilon_\infty}\sim(j_{ex}/j_{sc})^{2\varepsilon_c/X\varepsilon_\infty(X)} , \tag{4.63a}$$

which is equivalent to a nonlinear IV characteristic of the form

$$V\sim(j_{ex})^{a(X)} \tag{4.63b}$$

with

$$a(X)=\frac{2\varepsilon_c}{X\varepsilon_\infty(X)}+1 . \tag{4.63c}$$

The exponent $a(X)$ characterizing the nonlinear IV characteristics below T_c is another example of a Coulomb gas scaling function that should be "universal," provided the Ginzburg-Landau Coulomb gas description is applicable.

For $B_{ex}=0$ ($\omega^{rot}=0$) and $j_{ex}=0$ ($g_{ex}=0$) in the case of a superconducting (neutral superfluid) film below T_c, the macroscopic superfluid mass density is given by Eq. (4.43), i.e., $\rho_S/\rho=1/\varepsilon_\infty(X)$. Thus, $\rho_S(X)/\rho(X)$ is yet another example of a Coulomb gas scaling function.

The Coulomb gas scaling variables discussed above are listed in Table V, together with the corresponding variables for a superconducting film and a neutral superfluid film.

TABLE V. Coulomb gas scaling variables.

Coulomb gas variable	Superconducting film	Neutral superfluid film
T^{CG}/T_c^{CG}	X [see Eqs. (4.57) and (4.59)]	X
$r_0^2\langle\Delta n\rangle$	B_{ex}/B_{sc}, $B_{sc}=\varphi_0/(2\pi\xi^2)$	ω^{rot}/ω_{sc}, $\omega_{sc}=\frac{\hbar}{2m^*\xi^2}$
$r_0 D$	j_{ex}/j_{sc}, $j_{sc}=\frac{\varphi_0 c}{4\pi^2\Lambda\xi}$	g_{ex}/g_{sc}, $g_{sc}=\frac{\rho\hbar}{m^*\xi}$

E. *XY*-type models

In this section we discuss the connection between the phenomenological description of superfluid-superconducting films outlined in Secs. IV.A and IV.B and two-dimensional XY models.

The "standard" classical two-dimensional XY model is described by the Hamiltonian

$$H_{XY}=-J\sum_{\langle ij\rangle}\cos(\theta_i-\theta_j)\ , \tag{4.64}$$

where i and j denote the sites on a two-dimensional square lattice, the sum is over nearest-neighbor pairs of lattice points, J is a nearest-neighbor coupling constant, and θ_i is an angle associated with each lattice site. Alternatively one may associate a two-dimensional unit vector with each lattice site, and the interaction then involves the scalar product between vectors on neighboring sites. Obviously, the ground state corresponds to the situation when all such vectors point in the same direction. The statistical mechanics of the XY model is described by the partition function

$$Z_{XY}=\prod_i\int d\theta_i/(2\pi)e^{-H_{XY}/T}\ , \tag{4.65}$$

where T is the temperature. For a review of the two-dimensional classical XY model, see for example Suzuki (1979).

The two-dimensional XY model does not exhibit any true long-range order because the thermodynamic average $\langle\cos\theta_i\rangle$ is always zero (for $T\neq 0$). This lack of long-range order follows from the Mermin-Wagner theorem, which asserts that a broken continuous symmetry prevents long-range order in two dimensions (Mermin and Wagner, 1966). The two-dimensional XY model does, however, undergo a Kosterlitz-Thouless transition at a finite temperature T_c from a high-temperature phase (where the order-parameter correlation function has an exponential decay) to a low-temperature phase with "quasi" long-range order (where the order-parameter correlation has a power-law decay; Berezinskii, 1971; Kosterlitz and Thouless, 1972).

The two-dimensional XY model given by Eq. (4.64) can be viewed as an approximate microscopic description of a neutral two-dimensional superfluid (Matsubara and Matsuba, 1956, 1957; Hohenberg, 1971). This approximate microscopic description corresponds to associating an order parameter $\psi_j=|\psi|e^{i\theta_j}$ to each lattice site j where the magnitude of the order parameter is fixed and equal for all lattice points. The phase between neighboring lattice sites is coupled by the Hamiltonian

$$H=-\frac{1}{2}\left[\frac{\hbar}{m^*}\right]^2\sum_{\langle ij\rangle}(\psi_i^*\psi_j+\psi_j^*\psi_i)\ ,$$

which is identical to Eq. (4.64) with the identification $J=(\hbar/m^*)^2|\psi|^2$.

We now turn to the connection between the two-dimensional XY model and vortex fluctuations described by the two-dimensional Coulomb gas model. One way of establishing such a connection is the following (Kosterlitz and Thouless, 1973): XY-model configurations with small angular differences between neighboring sites, i.e., $|\theta_i-\theta_j|\ll 1$, are easily excited. One may hence expect that these are the important configurations for determining the thermodynamic properties. Provided all other configurations are neglected, H_{XY} can be approximated by expanding the cosine in Eq. (4.64),

$$H_{XY}\simeq -J\sum_{\langle ij\rangle}\left[1-\frac{(\theta_i-\theta_j)^2}{2}\right]\ . \tag{4.66}$$

The constant term in Eq. (4.66) may be absorbed into the definition of the ground-state energy and the remaining sum may be approximated by an integral (since $|\theta_i-\theta_j|\ll 1$), i.e.,

$$H_{XY}\simeq\frac{J}{2}\int d\mathbf{r}[\nabla\theta(\mathbf{r})]^2\ . \tag{4.67}$$

Finally, since we are expecting that the configurations with small angle differences between neighboring sites are the important ones, we may safely overlook the fact that in the XY model the angle $\theta+2n\pi$ is equivalent to θ and let θ take values on the complete interval $\{-\infty,\infty\}$. Now suppose that $\theta(\mathbf{r})$ [in Eq. (4.67)] is identified with the phase of the order parameter ψ for a neutral superfluid. It then follows that H_{XY} is equivalent to H_S given by Eq. (4.2) with the identification

$$J=\rho\left[\frac{\hbar}{m^*}\right]^2\ , \tag{4.68a}$$

since the superfluid mass flow $\mathbf{g}$ is given by

$$\mathbf{g}(\mathbf{r})=\frac{\rho\hbar}{m^*}\nabla\theta(\mathbf{r}) \tag{4.68b}$$

[compare Eq. (4.1)]. Consequently a vortex description similar to that given in Secs. IV.A.1 and IV.A.2 can be made. This chain of reasoning for the two-dimensional XY model thus leads to a description in terms of vortices and a two-dimensional Coulomb gas where the phase transition for the XY model corresponds to the charge-unbinding transition for the Coulomb gas (Kosterlitz and Thouless, 1972).

The smallest length scale on which a vortex can be defined for the two-dimensional XY model is obviously given by the lattice spacing. Figure 16 shows a vortex with vorticity $s=1$ on a square lattice with lattice spacing a. The angle θ_i associated with each lattice site is in the figure given by the angle between the arrows and a fixed direction.

The energy of a neutral configuration of N vortices defined on the lattice is approximately given by [Kosterlitz and Thouless, 1973; compare Eq. (2.16)]

$$H_N/(2\pi J)\simeq -\tfrac{1}{2}\sum_{k\neq 1} s_k s_1 \ln(r_{k1}/a)+N\frac{\pi}{4}\ , \tag{4.69}$$

where the sum is over all vortex pairs. A comparison with Eq. (2.16) shows that, in this "lattice" approximation of the Coulomb gas, E_c is determined by the energy needed to create a neutral vortex pair with the two vortices on nearest-neighbor sites, which gives $E_c\simeq\pi/4$ (Kosterlitz and Thouless, 1973). The phase-space division ζ for a vortex on the lattice is given by $\zeta=a^2$.

Let us now discuss some aspects of the link between the XY model and the phenomenological type of Coulomb gas description given in Secs. IV.A.1, IV.A.2, IV.D.1, and IV.D.2.

First consider the approximation of the XY Hamiltonian given by Eq. (4.67). The corresponding approximation

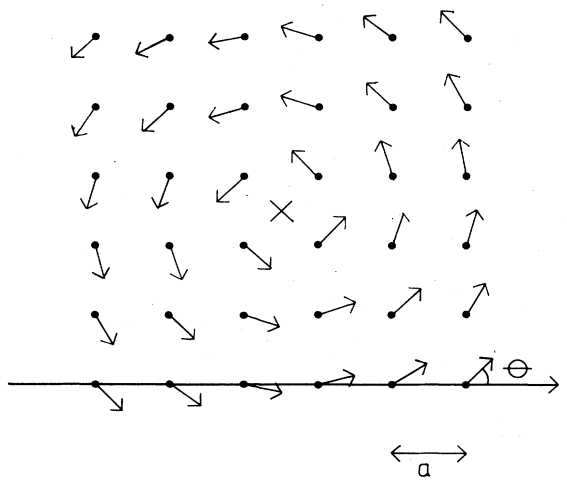

FIG. 16. A configuration for the XY model corresponding to a vortex with vorticity $s=1$. The figure shows a square lattice with lattice spacing a (dots). An angle θ_i is associated with each lattice site. This angle is represented as the angle between the arrows and a fixed direction (horizontal axis in the figure). The center of the vortex is denoted by a cross.

of the thermodynamics of the XY model can be improved by replacing J in Eqs. (4.66) and (4.67) with a "renormalized" $J_R(T)$. One method of achieving such an improvement is through a minimization of the free energy within a self-consistent harmonic approximation, which gives (Lozovik and Akapov, 1981)

$$J_R(T)\simeq J\left[1-\frac{T}{T_{c0}}\right]\ , \tag{4.70}$$

with $T_{c0}=nJ$ where n is the number of nearest neighbors. $J_R(T)$ is related to the "bare" superfluid density $\rho(T)$ by $J_R(T)=\rho(T)(\hbar/m^*)^2$ [compare Eqs. (4.68)].

An infinitesimal constant macroscopic superfluid mass flow $\delta\mathbf{g}_{\rm tot}$ is related to an imposed infinitesimal constant phase gradient [compare Eq. (4.68b)], i.e., if the imposed constant gradient of the phase θ is denoted by $\delta\mathbf{h}_{\rm ex}$, then the relation is $\delta g_{\rm tot}=(\rho_S\hbar/m^*)\delta h_{\rm ex}$ where ρ_S is the macroscopic superfluid density. Consequently the correspondence of Eqs. (4.9) is

$$\Delta F=\frac{\Omega\rho_S}{2}\left[\frac{\hbar}{m^*}\right]^2(\delta h_{\rm ex})^2 \tag{4.71a}$$

or alternatively

$$\rho_S\left[\frac{\hbar}{m^*}\right]^2=\gamma\ , \tag{4.71b}$$

where

$$\gamma=\frac{\partial^2 F_{XY}}{\Omega\partial h_{\rm ex}^2}\bigg|_{\delta h_{\rm ex}=0}\ . \tag{4.71c}$$

The quantity γ is called the helicity modulus, and the relation between γ and ρ_S [Eq. (4.71b)] was established by Fisher *et al.* (1973). For the XY model on a square lattice, γ is given by (Teitel and Jayaprakash, 1983)

$$\gamma=-\frac{a^2}{2\Omega}\langle H_{XY}\rangle-\frac{Ja^2}{T\Omega}\left\langle\left[\sum_{\langle ij\rangle}\sin(\theta_i-\theta_j)\hat{\mathbf{e}}_{ij}\cdot\hat{\mathbf{x}}\right]^2\right\rangle\ , \tag{4.72}$$

where $\hat{\mathbf{e}}_{ij}$ ($\hat{\mathbf{x}}$) is unit vector pointing from the lattice point i to the nearest-neighbor point j (pointing in an arbitrary fixed direction in the plane of the lattice) and $\langle\ \rangle$ denotes the thermodynamic average. The universal jump prediction for the helicity modulus of the XY model may be obtained by multiplying Eq. (4.44) by $(\hbar/m^*)^2$ and using Eq. (4.71b), which gives

$$\lim_{T\to T_c^-}\frac{\gamma(T)}{k_B T}=\frac{2}{\pi}\ . \tag{4.73}$$

Figure 17 shows γ for the XY model as obtained from typical Monte Carlo simulations. The crossing between the $\gamma(T)$ curve and the line $\gamma=(2/\pi)TJ$ determines the critical temperatures T_c. Since simulations by necessity are done on finite lattices, $\gamma(T)$ has only a rapid decrease at T_c and not a discontinuous jump to zero. Also plotted in the figure is the specific-heat peak, which appears at a

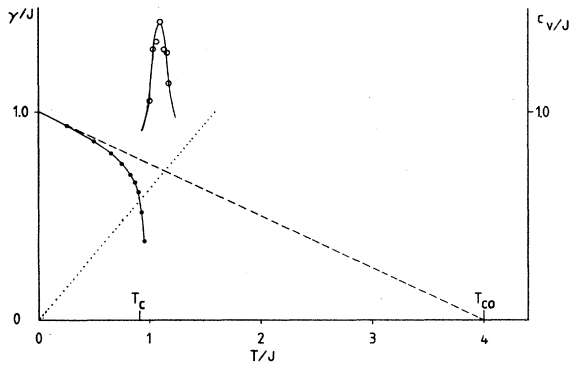

FIG. 17. The helicity modulus and the specific-heat peak for the XY model on a square lattice, as obtained by Monte Carlo simulations. Solid dots are from a Monte Carlo simulation of the helicity modulus γ (63×64 lattice, from Minnhagen and Nylén, 1985b). Dashed line is the exact low-temperature asymptote $\gamma/J=1-T/T_{c0}$, where $T_{c0}=4J$. Dotted line is $\gamma/J=2T/\pi$. Crossing between the solid curve through solid dots and the dotted line defines the Kosterlitz-Thouless temperature T_c. Open circles are a Monte Carlo simulation of the peak in the specific heat C_V (40×40 lattice, from an investigation of the specific-heat peak by van Himbergen and Chakravarty, 1981). The specific-heat peak occurs above T_c and does not reflect any critical properties.

temperature higher than T_c. The specific-heat peak does not reflect any nonanalytical critical properties, in contrast for example to the case of first- and second-order phase transitions (Kosterlitz, 1974; Berker and Nelson, 1979).

The connection between the XY model and the Coulomb gas can also be obtained in a more systematic way through a duality transformation (José *et al.*, 1977; Savit, 1978). Such an analysis shows that Eq. (4.70) is exact in the small-temperature limit (Ohta and Jasnow, 1979). In analogy with the model of a superfluid described in Sec. IV.A.1, the helicity modulus excluding vortex fluctuations is given by $\gamma=\rho(T)(\hbar/m^*)^2$, where $\rho(T)$ is the "bare" superfluid density [compare Eq. (4.71b) and the paragraph just below Eq. (4.11)]. Since the vortices require a finite energy to be excited [see Eq. (4.69)], it follows that the "bare" helicity modulus is exactly given by $\gamma=J(1-T/T_{c0})$ with $T_{c0}=4J$ in the limit of small T (see Fig. 17). Furthermore, this result is independent of the length scale on which the vortices are defined.

On the other hand, for temperatures larger than $T=0$, the value of γ excluding vortex fluctuations obviously depends in general on the length scale on which the vortices are defined [see the paragraph just below Eq. (4.2)]. On each length scale one can consider γ (or, more generally, any thermodynamic average of the XY model that translates into a thermodynamic average in the Coulomb gas analogy) as having a "bare" ("full") value corresponding to excluding (including) vortex fluctuations. The "bare" value can be expressed in terms of "bare" Coulomb gas parameters. The "bare" Coulomb gas parameters on different length scales are related through the renormalization equations of the Coulomb gas model described in Sec. III.

The basis for the Ginzburg-Landau Coulomb gas described in Sec. IV.D is that the free energy can be described by a Ginzburg-Landau approximation in terms of a complex order parameter $\psi(\mathbf{r})=|\psi(\mathbf{r})|e^{i\theta(\mathbf{r})}$ [see Eq. (4.49)]. For such a description to apply to the XY model, it is required that the free energy for the XY model approximately "renormalize" into a Ginzburg-Landau free energy on some length scale larger than the lattice spacing. An indication of how such a transformation from an XY model to a Ginzburg-Landau description can be achieved (using a Hubbard-Stratanovich transformation) is given by Doniach (1984).

As mentioned above, the "standard" two-dimensional XY model [see Eq. (4.64)] can be viewed as an approximate microscopic description of a neutral two-dimensional superfluid. A corresponding lattice model for a two-dimensional charged superfluid is given by (Lozovik and Agapov, 1980)

$$H_{XY}=-J\sum_{\langle ij\rangle}\cos(\theta_i-\theta_j-A_{ij})+\int d\mathbf{x}\,B^2(\mathbf{x})/(8\pi)\ , \tag{4.74a}$$

where $\mathbf{x}=(\mathbf{r},x_3)$, the two-dimensional lattice is in the $x_3=0$ plane, A_{ij} is related to the vector potential $\mathbf{A}(\mathbf{x})$ by the line integral along the straight line connecting site i with site j, i.e.,

$$A_{ij}=\frac{e^*}{\hbar c}\int_{\mathbf{r}_i}^{\mathbf{r}_j}d\mathbf{r}\cdot\mathbf{A}(\mathbf{x})\ , \tag{4.74b}$$

the $\mathrm{div}(\mathbf{A})=0$ gauge is chosen, and the induced magnetic field $\mathbf{B}$ is given by $\mathbf{B}=\mathrm{curl}(\mathbf{A})$. The partition function involves fluctuations of the vector potential in addition to the fluctuations of the phase angles, i.e.,

$$Z_{XY}=\int d\,\mathbf{A}(\mathbf{x})\delta(\mathrm{div}(\mathbf{A}))\prod_i\int d\theta_i/(2\pi)e^{-H_{XY}/T}\ . \tag{4.75}$$

This may be viewed as an approximate model of a granular superconducting film or a two-dimensional array of weakly coupled Josephson junctions, provided charging fluctuations are negligible (Lozovik and Akapov, 1980). Note that because of the coupling to the electromagnetic field it is a three-dimensional model. In analogy with the connection between the "standard" XY model and the model of a neutral superfluid described in Sec. IV.A, one may expect that, on a sufficiently large length scale compared to the lattice spacing, this lattice model of a superconductor will "renormalize" into the description of a superconducting film given in Sec. IV.B. Likewise one may expect that on a sufficiently large length scale a Ginzburg-Landau approximation may be adequate (see Sec. IV.D).

V. EXAMPLES OF EXPERIMENTS REFLECTING STATIC COULOMB GAS PROPERTIES

The aim of this section is to illustrate the connection between the two-dimensional Coulomb gas described in the previous sections and experiments on ^{4}He films and superconducting films. To this end a few examples from the large body of experimental work to date are selected and discussed. They are chosen with the criteria of reflecting static Coulomb gas properties as well as possible (as opposed to experiments where the dynamics of the vortices is crucial for the interpretation). The choice of examples is subjective.

The reader interested in a more complete description of the experimental work is referred to review articles by Hebard and Fiory (1982), Mooij (1983), and Mooij (1984) in the case of superconducting films, and Glaberson and Donelly (1986) and Kotsubo and Williams (1986) in the case of helium films, which cover at least part of the large amount of experimental work in this area.

Section V.A treats two special (static) Coulomb gas aspects of the superfluid density for ^{4}He films as measured by the torsion pendulum experiment devised by Bishop and Reppy (1978). The first aspect is the size of the jump of the superfluid density at the superfluid transition (Sec. V.A.1), while the second is the Coulomb gas scaling of the superfluid density (Sec. V.A.2). Section V.B treats a variety of (static) Coulomb gas aspects of superconducting films. Section V.B.1 deals with the "universal" resistance curve and its connection to the Ginzburg-Landau Coulomb gas. The nonlinear IV characteristics and the extracted Coulomb gas dielectric constant are discussed in Sec. V.B.2. Finally Sec. V.B.3 describes the magnetic field scaling of the resistance.

A. ^{4}He films

The torsion pendulum experiment devised by Bishop and Reppy (1978) gives a particularly direct measurement of the superfluid density and is for this reason chosen to illustrate the connection to (static) Coulomb gas properties. This experiment is set up in such a way that the helium is absorbed on a Mylar substrate. The substrate is rolled in a jelly roll fashion and this jelly roll is part of the pendulum. The superfluid part of the helium does not, however, contribute to the oscillating mass (unless the superflow relative to the substrate is hindered by geometrical constraints). Thus the frequency of the pendulum is in principle given by

$$\omega=\left[\frac{K}{M-\Omega\rho_S}\right]^{1/2}\simeq\sqrt{K/M}+\sqrt{K/M}\,\frac{\Omega}{2M}\rho_S\ , \qquad (5.1)$$

where K is the spring constant and M is the total mass (including the mass of the helium). Consequently the period shift of the oscillator, $\Delta P\simeq\sqrt{M/K}-\omega^{-1}$, is proportional to the superfluid density ρ_S. For a detailed description of the experiment see Bishop and Reppy (1980). A complete analysis of the experiment also involves the vortex dynamics that cause dissipation. This will be discussed in Sec. VI.

1. The universal superfluid density jump

Figure 18 shows the period shift ΔP as a function of the temperature for six different ^{4}He-^{3}He mixture films obtained by McQueeney *et al.* (1984). The (static) Coulomb gas prediction is that the quantity $1/T\varepsilon_\infty$ drops discontinuously from 4 to 0 at the Kosterlitz-Thouless transition. This in turn is equivalent to the universal jump prediction for the superfluid density [Nelson and Kosterlitz, 1977; see Eq. (4.44)]. The solid line in Fig. 18 corresponds to this universal jump prediction. Consequently the period shift should drop discontinuously to zero at the crossing with the respective period-shift curves. As seen in the figure, this prediction is realized rather well in that the period shift in all cases starts to drop rapidly to zero at these crossings.

The fact that the measured period shift does not drop discontinuously to zero can be attributed to the diffusive motion of the vortices, caused by the fact that the experiment is performed at a finite frequency (Ambegaokar *et al.*, 1978, 1980); this will be discussed further in Sec. VI. This effect can be qualitatively inferred without any detailed dynamical assumptions in the following way (compare Sec. VI.A.1): Denote the frequency of the torsion pendulum by ω and the diffusion constant for the vortex motion by $\mathcal{D}$. The length $\sqrt{\mathcal{D}/\omega}$ is the characteristic length for vortex diffusion, which in turn means that only vortex pairs with separation much less than $\sqrt{\mathcal{D}/\omega}$ will respond adiabatically to the oscillation (Ambegaokar and Teitel, 1979). This means that the effective Coulomb gas length scale probed by the experiment is of the order of $\lambda_{\rm exp}\sim\sqrt{\mathcal{D}/\omega}$. Consequently, from Eq. (4.48) and the connection given by Eq. (5.1), one infers that ΔP should drop *continuously* to zero with the high-

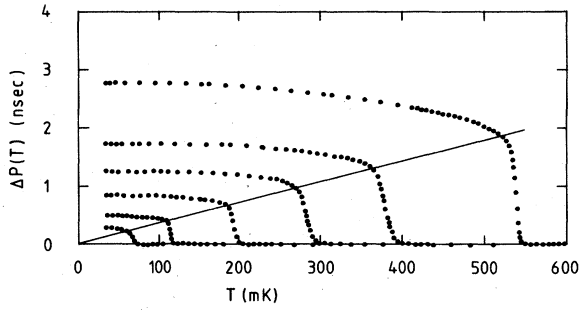

FIG. 18. The period shift ΔP as a function of temperature T for six different ^{4}He-^{3}He-mixture films. ΔP is proportional to the superfluid density. The straight line corresponds to the universal jump prediction. The superfluid density drops rapidly to zero at the crossing between the measured curves and the straight line, in accordance with the universal jump prediction (from McQueeney *et al.*, 1984).

temperature tail roughly proportional to $(\lambda_F/\lambda_{\rm exp})^2$, where λ_F is the Coulomb gas screening length due to free vortices.

2. Scaling properties of the superfluid density

The period shifts measured in the torsion pendulum experiment show scaling properties (Agnolet, 1983; Agnolet *et al.*, 1984b). The Coulomb gas scaling ideas discussed in Sec. IV.D.2 offer a possible explanation of these observed scaling properties (Minnhagen and Nylén, 1985b). Consequently these scaling properties may be viewed as a link between the two-dimensional Coulomb gas and the measurements. The basic assumption invoked in order to obtain this link is that the vortex fluctuations of the various samples are described by the same Coulomb gas theory.

Figure 19 shows a prototypical measurement. The solid curve is the measured period shift ΔP, which is proportional to the superfluid density ρ_S [see connection given by Eq. (5.1)]. The ΔP curve is, to a good approximation, a horizontal line for lower temperatures (dashed line in Fig. 19). This horizontal line is proportional to the bare superfluid density ρ, since the vortices drop out for low enough temperatures due to their finite excitation energy. It is further assumed that on some length scale this constant ρ describes the bare superfluid density for all temperatures. This means that the Coulomb gas scaling variable X reduces to $X=T/T_c$, since a constant ρ corresponds to $T_{c0}=\infty$ [compare Eqs. (4.59)]. The Coulomb gas dielectric constant is given by the ratio ρ/ρ_S [see Eq. (4.43)]. In Fig. 19 this is just the ratio between the ΔP values given by the dashed line and the solid curve. The dotted line in Fig. 19 is the universal jump prediction [compare Eq. (4.44) and Fig. 18]. The crossing between this line and the solid curve identifies T_c. The Coulomb gas dielectric constant $\varepsilon_\infty(X)$ obtained in this way for three different samples is shown in Fig. 20 [the measurements are taken from McQueeney *et al.* (1984) and Agnolet *et al.* (1984a); see Minnhagen and Nylén (1985b) for a more detailed description]. As is apparent from Fig. 20, the Coulomb gas scaling for $\varepsilon_\infty(X)$ is well obeyed.

Coulomb gas scaling may also be invoked to test the link between the helium films and the XY model described in Sec. IV.E (Minnhagen and Nylén, 1985b). Figure 17 shows a Monte Carlo simulation of the helicity modulus γ, which is proportional to ρ_S [see Eq. (4.71b)]. The bare superfluid density ρ is in this case temperature dependent and is, to lowest approximation, proportional to the dashed line in Fig. 17, while T_{c0} is equal to $4J$. The Coulomb gas dielectric constant is the ratio of the γ values given by the dashed line and the solid curve in Fig. 17. Figure 21 shows a comparison between the $\varepsilon_\infty(X)$ obtained for the XY model on a square lattice (solid line) and that obtained from measurements on helium films (dashed line—the average of the measurements given in Fig. 20). The agreement indicates a strong connection between the vortex fluctuations described by the XY model and the vortex fluctuations observed for helium films.

The dotted line is a calculation based directly on the "lattice" Coulomb gas approximation for the XY model given by Eq. (4.69) and the renormalization equations [Eqs. (3.14)]. The dotted curve is produced by integrating the renormalization equations from the initial length a (the lattice constant) and the initial values $z_i=2e^{-\pi/(4XT_c)}$ and $T_i=\varepsilon_i T_c X$, where the factor 2 is an entropy factor reflecting the fact that there are precisely four possible positions for one particle in a neutral pair with separation

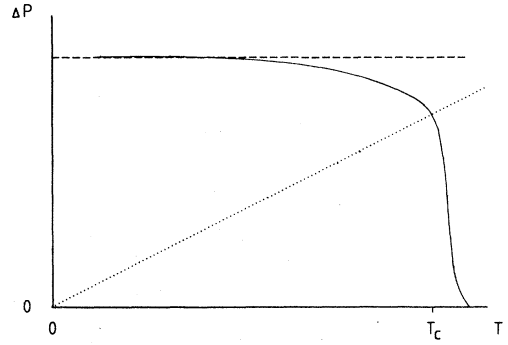

FIG. 19. Construction of an approximate Coulomb gas interpretation. The period shift ΔP is measured as a function of T. ΔP is proportional to the superfluid density. The dotted line corresponds to the universal jump prediction. The crossing between the solid curve and the dotted line gives an estimate of the Kosterlitz-Thouless temperature T_c. The dashed horizontal line is proportional to the bare superfluid density. This corresponds to $T_{c0}=\infty$ and the Coulomb gas scaling variable $X=T/T_{c0}$. The ratio between the values given by the dashed line and the solid curve gives the Coulomb gas dielectric constant ε_∞.

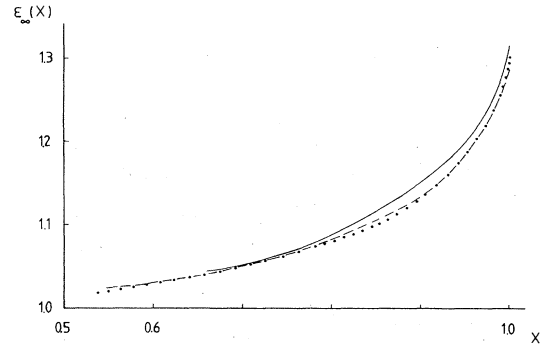

FIG. 20. The Coulomb gas dielectric constant $\varepsilon_\infty(X)$ extracted from the torsion pendulum experiment in the way described in connection with Fig. 19. The solid, dashed, and dotted curves represent three different samples (from Minnhagen and Nylén, 1985b, where more details are given; the T_c values for the three samples corresponding to the solid, dashed, and dotted lines are $T_c=365$, 211, and 165 mK, respectively, and $T_{c0}=\infty$ in all cases; the actual measurements are taken from McQueeney *et al.*, 1984, and Agnolet *et al.*, 1984a). The $\varepsilon_\infty(X)$ extracted from the three samples are in good agreement. This may be interpreted as an indication of Coulomb gas scaling.

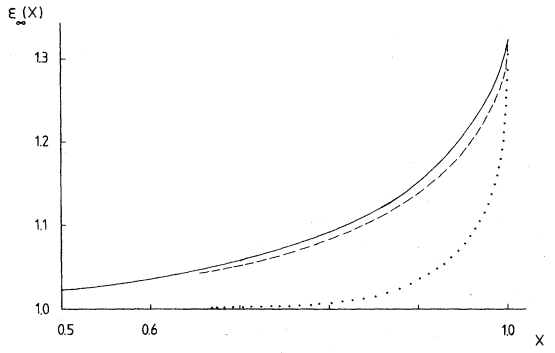

FIG. 21. The Coulomb gas dielectric constant $\varepsilon_\infty(X)$ extracted from the XY model on a square lattice (solid curve) and from helium films (dashed curve = average of the measurements given in Fig. 20). The agreement indicates a strong connection between vortex fluctuations for the XY model and helium films. The dotted curve is based on the "lattice" Coulomb gas approximation of the XY model together with the Coulomb gas renormalization equations (see text).

a relative to the other in the case of a square lattice; $T_c = 1/4\varepsilon_c \simeq 0.19\varepsilon_i$ is the critical Coulomb gas temperature corresponding to the XY model on a square lattice. The initial value ε_i can be estimated by $\varepsilon_i = (1 - 2\pi\chi_0)^{-1} \simeq 1 + 2\pi\chi_0$, where χ_0 is the polarizability due to neutral pairs with separation a. The density of such pairs can be estimated by z_i^2/a^2 and the polarizability of such a pair is $a^2/(2T)$ [see discussion below Eq. (3.4)], so that $\varepsilon_i \simeq 1 + \pi z_i^2/(T_c X)$. It is obvious from Fig. 21 that the "lattice" Coulomb gas approximation given by Eq. (4.69) combined with the renormalizations equations in the above fashion does not give a good description of the $\varepsilon_\infty(X)$ extracted from the Monte Carlo data for the XY model nor of the $\varepsilon_\infty(X)$ determined from experiments on helium films. This indicates that the "lattice" Coulomb gas is a rather crude approximation.

In summary, the situation for the helium films is the following: The Coulomb gas universal jump prediction for the superfluid density and the Coulomb gas scaling seem to be borne out by experiments. The vortex fluctuations described by the XY model appear to be strongly related to the vortex fluctuations for ^{4}He films. On the other hand, the "lattice" Coulomb gas model by itself does not give an adequate description.

B. Superconducting films

The Coulomb gas properties are strongly reflected in experimental data from superconducting films, as will be shown below. In fact the connection to the (static) Coulomb gas appears at present to be more firmly established for superconducting films than for helium films.

The experiments chosen below to illustrate the connection between Coulomb gas properties and superconducting films are measurements of IV (current-voltage) characteristics.

1. The "universal" resistance curve

According to the Coulomb gas scaling ideas discussed in Sec. IV.D.2, the resistance ratio R/R_N should be a "universal" (i.e., sample-independent) function of the scaling variable X given by Eq. (4.59b) in the absence of an external perpendicular magnetic field. Figure 22 shows that this scaling property is indeed borne out by experiment. The figure demonstrates how data from five different samples collapse onto a single "universal" curve when plotted against the scaling variable X (the figure is taken from Minnhagen, 1983a, where identification of the measurements and further details can be found).

The "universal" resistance curve constructed from available data is given in Fig. 23 (from Minnhagen, 1984b). An obvious question concerning this resistance curve is to what extent its functional form can be tied to the theory of the two-dimensional Coulomb gas. A commonly used approach is to try to relate the temperature dependence of the resistance curve to the Kosterlitz renormalization equations [Eqs. (3.16)]. This possible relationship was first pointed out by Halperin and Nelson (1979). The connection is basically the following: From Eqs. (4.58) and (2.21b) one obtains a relation between the Coulomb gas screening length λ and the resistance ratio, which by use of the definition of X [Eq. (4.59a)] and the relation $4T_c^{CG}\varepsilon_c = 1$ can be turned into

$$R/R_N = \tilde{\varepsilon}(X) X \xi^2/[4\varepsilon_c \lambda^2(X)] \; . \tag{5.2}$$

Close to the Kosterlitz-Thouless transition the screening length λ can be approximated by Eq. (3.28), which leads to

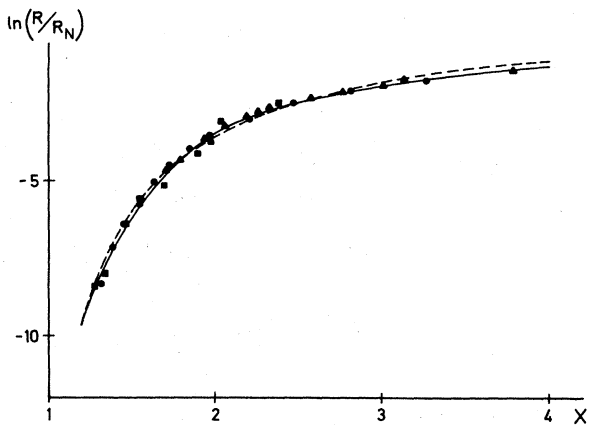

FIG. 22. The measured resistance ratio R/R_N plotted against the Coulomb gas scaling variable $X = T(T_{c0} - T_c)/[T_c(T_{c0} - T)]$ for five different superconducting films. The measurements are represented by squares, dots, triangles, solid curve, and dashed curve (from Minnhagen, 1983a, where an identification of the measurements are given, together with the values of T_c and T_{c0} and further details). The resistance ratios plotted against X are apparently well described by a single function.

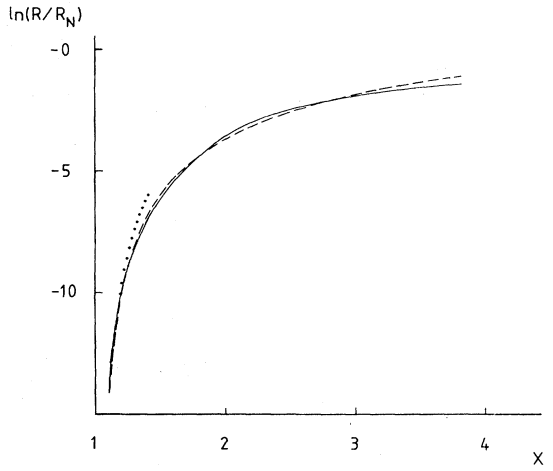

FIG. 23. The "universal" resistance curve constructed from data (solid curve, from Minnhagen, 1984b). The dashed curve is a fit to the functional form $R/R_N = C_0 X e^{-C_1/\sqrt{X-1}}$ using C_0 and C_1 as free parameters [compare Eq. (5.4a); the dashed curve corresponds to $C_0=1.7$ and $C_1=4.9$]. The agreement is good, but the physical relevance is questionable. The dotted curve is a fit to the same functional form but with C_1 estimated *ab initio* ($C_1=5.3$) and C_0 as a free parameter ($C_0=7.5$). The agreement is good in the region where the functional form can be motivated.

$$R/R_N = C_0 X e^{-C_1/\sqrt{X-1}} \quad \text{for } 0 < X-1 \ll 1 , \tag{5.3}$$

where C_0 and C_1 are two constants. The constant C_1 can be related to the initial values $\varepsilon_i(T_c)$ and $z_i(T_c)$ of the renormalization-group equations at T_c [compare Eq. (3.28)], i.e.,

$$C_1 = \left[\frac{\pi \varepsilon_i(T_c)}{4 z_i(T_c) \varepsilon_c} \right]^{1/2} . \tag{5.4}$$

The dashed curve in Fig. 23 represents an attempt to fit the experimentally determined resistance curve with the functional form given by Eq. (5.3) using C_0 and C_1 as fitting parameters. As can be seen in the figure, the temperature dependence of the resistance can be very well represented by the functional form given by Eq. (5.3).

This sort of agreement between the temperature dependence of the resistance and the functional form determined from the Kosterlitz RG equations is frequently taken as an indication that the measured resistance is indeed caused by vortex fluctuations (e.g., Resnick *et al.*, 1981; Abraham *et al.*, 1982; Voss and Webb, 1982; Fiory *et al.*, 1983; Hebard and Fiory, 1983; Kadin *et al.*, 1983). However, it has been pointed out that this conclusion may be too rash (Bancel and Gray, 1981; Minnhagen, 1981b; Minnhagen, 1983b; Gray *et al.*, 1985). First of all the theoretical prediction presumes that $0 < X-1 \ll 1$. Consequently the significance of the agreement outside this region is questionable. Second, the measured resistance curve has so little structure that a reasonable fit using the functional form given by Eq. (5.3) with two or more adjustable parameters is almost guaranteed, without any particular physical significance can be ascribed to it.

One way of making a more severe test of the connection between the functional form given by Eq. (5.3) for $0 < X-1 \ll 1$ and the resistance curve would be to reduce the number of free parameters. The dotted line in Fig. 23 represents such an attempt. In this case the constant C_1 has been estimated as $C_1 \simeq 5.3$ by using Eq. (5.4) in the following way (Minnhagen, 1986): The value $\varepsilon_c \simeq 1.65$ is determined from experiments (Minnhagen, 1983b). The value $z_i(T_c)$ is then estimated from a Monte Carlo simulation of the Ginzburg-Landau Coulomb gas (Minnhagen and Weber, 1985; see below). Finally, the requirement that this be consistent with the renormalization equations gives $\varepsilon_i(T_c) \simeq 1.39$, and $C_1 \simeq 5.3$ follows. As is apparent from a comparison of the dotted and solid curves in Fig. 23, the functional form given by Eq. (5.3) agrees well with the temperature dependence of the resistance in the interval $1.1 < X < 1.2$ when the constant $C_1 \simeq 5.3$ is given *ab initio* and only the parameter c_0 is varied. This may be taken as an indication of agreement between theory and experiment close to the Kosterlitz-Thouless transition, although a firmer conclusion as regards the functional form given by Eq. (5.3) would require that the constant C_0 also be estimated *ab initio*.

The obvious weakness of the functional form of R/R_N derived from the Kosterlitz RG equations is that it can only be motivated for $0 < X-1 \ll 1$, whereas the overwhelming part of the data is from outside this region. Furthermore, close enough to $X=1$ the theoretical functional form derived from the Kosterlitz RG equations breaks down in the case of a superconductor, due to the fact that λ_c is finite (compare Fig. 2 and Table III). A possible way of overcoming this difficulty in the comparison between theory and experiment is to try to obtain the quantities on the right-hand side of Eq. (5.2) outside the region $0 < X-1 \ll 1$ by other means. One such means is to resort to Monte Carlo simulations for the two-dimensional Coulomb gas.

An example of an attempt in this direction is the following (Minnhagen and Weber, 1985): The Ginzburg-Landau Coulomb gas (see Sec. IV.D) is transformed into a nonlocal sine-Gordon theory (see Sec. II.C), letting E_c and $\tilde{c}$ be free parameters (compare Table IV). This nonlocal sine-Gordon theory is discretized into a lattice field theory, and the Fourier transform of the Green's function at $k=0$ [see Eq. (2.60)] is generated by a Monte Carlo simulation technique. The connection to the resistance in this case takes the form $R/R_N = X\hat{G}^{-1}(k=0)/(4\varepsilon_c)$ [compare Eqs. (2.61), (2.64), and (5.2)]. The value of ε_c is estimated to be $\varepsilon_c \simeq 1.65$ from experimental data for nonlinear IV characteristics (Minnhagen, 1983b; see Sec. V.B.2). The free parameter E_c in the Monte Carlo simulation is varied in order to obtain the best fit to the experimental R/R_N curve (only the portion of the curve $X > 1.8$ is used, due to limitations in the simulation convergence). A value $E_c \simeq 0.36$, close to the Ginzburg-Landau Coulomb gas value $E_c = 0.39$, has been obtained in this way (Minnhagen and Weber, 1985). This was interpreted

as an indication that the "universal" resistance curve extracted from experiments is closely related to the Ginzburg-Landau Coulomb gas model (Minnhagen and Weber, 1985).

As a by-product of this Monte Carlo simulation, the value $\tilde{c}\simeq 13$ was obtained (compare Table IV). This value, together with the Ginzburg-Landau Coulomb gas value $E_c=0.390$ (see Table IV) and the experimentally determined $\varepsilon_c\simeq 1.65$ (Minnhagen, 1983b), gives the initial value $z_i(T_c)\simeq 0.023$ used above in order to estimate C_1 given by Eq. (5.4).

Thus, in summary, the situation for the "universal" resistance curve seems to be the following: The Ginzburg-Landau Coulomb gas prediction, that the experimental data for the resistance of thin superconducting films should fall on a single curve when plotted against the Coulomb gas scaling variable X, appears to be confirmed (Minnhagen, 1981b; Resnick *et al.*, 1981; Minnhagen, 1983a, 1983b; Kihmi *et al.*, 1984; Minnhagen, 1984b; Lee, 1985). The functional form of the resistance predicted by the Kosterlitz RG equations is consistent with the data (e.g., Bancel and Gray, 1981; Resnick *et al.*, 1981; Wolf *et al.*, 1981; Abraham *et al.*, 1982; Voss and Webb, 1982; Fiory *et al.*, 1983; Hebard and Fiory, 1983; Kadin *et al.*, 1983; Svistunov *et al.*, 1984), *but* the physical significance is questionable due to too many free parameters combined with the fact that most of the resistance data are from outside the region where the theoretical prediction can be motivated. Reducing the number of free parameters to one, as described above, to some extent overrules this objection and does give a result that agrees in the region where the theoretical prediction can be motivated. Hence it may be taken as an indication of agreement between Coulomb gas theory and experiment. A stronger connection is indicated by Monte Carlo simulations which suggest a close relation between the Ginzburg-Landau Coulomb gas model and the "universal" resistance curve (Minnhagen and Weber, 1985).

2. Nonlinear IV characteristics

According to the Coulomb gas model there are no free vortices present below T_c and hence no flux-flow resistance. However, free vortices should, according to the Coulomb gas model, be generated below T_c provided a finite current j_{ex} is imposed across the superconducting film. The Coulomb gas prediction for the flux-flow resistance generated below T_c in this way is equivalent to a nonlinear IV characteristic of the form $V\sim(j_{ex})^{a(X)}$, where $a(X)=1+2\varepsilon_c/[X\varepsilon_\infty(X)]$ [see Sec. IV.D.2 in connection with Eqs. (4.63)]. Measurements of this nonlinear IV characteristic below T_c have been performed by Epstein *et al.* (1981, 1982), Resnick *et al.* (1981), Abraham *et al.* (1982), Fiory *et al.* (1983), Hebard and Fiory (1983), Kadin *et al.* (1983), and Lee (1985).

Precisely at T_c the Coulomb gas prediction for the exponent is $a(X=1)=3$ [compare Eq. (4.63c)]. This prediction is borne out by experiments, as has been especially clearly demonstrated by Hebard and Fiory (1983) and Fiory *et al.* (1983) on indium/indium oxide composite films.

In the experiments by Fiory *et al.* (1983) it was also shown that the Coulomb gas interpretation of the nonlinear IV characteristics agreed well with the Coulomb gas interpretation of measurements of the complex impedance very close to T_c. However, further below T_c these two types of measurements diverged (Fiory *et al.*, 1983), suggesting that other effects come into play at lower temperatures, possibly vortex pinning. Deviations from the pure Coulomb gas interpretation below T_c due to vortex pinning have also been suggested by Lee (1985) on the basis of experiments of flux-flow resistance in a perpendicular magnetic field for the same type of indium/indium oxide samples as used by Fiory *et al.* (1983). This type of additional complication must be kept in mind when trying to interpret experiments by the Coulomb gas model, as has been emphasized by Gray *et al.* (1985). On the other hand, the importance of such additional complications seems to vary among different types of samples. In order to test the Coulomb gas predictions as directly as possible, one obviously needs a sample for which these additional complications are suppressed as much as possible. Possible candidates in this respect are the Hg-Xe alloy films measured by Epstein *et al.* (1982) and Kadin *et al.* (1983). These samples may, *ipso facto,* be interpreted by the Coulomb gas model in a particularly consistent way, including temperatures extending somewhat below the Kosterlitz-Thouless temperature (Minnhagen, 1983b). In this section these "favorable" measurements have been chosen in order to illustrate various aspects of the connection to the Coulomb gas model.

In Figs. 24(a) and 24(b), results from measurements of the nonlinear IV characteristics for two Hg-Xe alloy films are plotted (samples no. 4 and 5 from Kadin *et al.*, 1983). The measured exponent $a(T)$ is plotted as $[a(T)-1]T$ against T [dots in Figs. 24(a) and 24(b), from Minnhagen, 1983b]. According to the Coulomb gas analogy, $a(X)$ is given by Eq. (4.63c). Provided that a Ginzburg-Landau approximation is adequate and that X is given by Eq. (4.59b), one obtains

$$[a(T)-1]T\sim(T_{c0}-T)/\varepsilon_\infty(T)\ , \tag{5.5}$$

where T_{c0} is the Ginzburg-Landau temperature and $\varepsilon_\infty(T)$ describes the polarization due to bound vortex pairs. As is explained in connection with Fig. 3, $\varepsilon_\infty(T)$ is approximately 1 when T is small enough. Hence one expects that $[a(T)-1]\sim(T_{c0}-T)$ for lower temperatures. As can be seen in Figs. 24(a) and 24(b), this expectation of a linear temperature dependence is realized in a narrow temperature region (the solid lines). On the other hand, for too low temperatures this linearity appears to break down. Note that according to this Coulomb gas interpretation the crossing between the solid line and the T axis defines T_{c0}. The dashed line in Figs. 24(a) and 24(b) is

$2T$ plotted against T. Since $a(X)=3$ at the Kosterlitz-Thouless transition, it follows that the crossing between the dashed line and the measured $[a(T)-1]T$ points defines the Kosterlitz-Thouless temperature T_c.

Taken by itself the Coulomb gas interpretation of the nonlinear IV characteristics given by figures like 24(a) and 24(b) is consistent but, of course, hardly compelling. However, the evidence in favor of this interpretation increases considerably if, when taken together with other data, it forms part of a consistent picture. As an example of such a consistency test, Fig. 25 shows the flux-flow resistance for the sample corresponding to Fig. 24(a) plotted against $X=T(T_{c0}-T_c)/[T(T_{c0}-T)]$ (dots in Fig. 25), with T_{c0} and T_c obtained from the Coulomb gas interpretation of the nonlinear IV characteristics. The solid curve is the "universal" resistance curve extracted from entirely different types of samples. As can be seen in Fig. 25, the dots fall on the "universal" curve. This means that for this Hg-Xe alloy sample the Coulomb gas interpretation of the nonlinear IV characteristics below T_c agrees perfectly with the Coulomb gas interpretation of the flux-flow resistance above T_c (Minnhagen, 1983b).

Another possible consistency test is to extract the Coulomb gas dielectric constant $\varepsilon_\infty(X)$ from Figs. 24(a) and 24(b). For a given X, $\varepsilon_\infty(X)$ is the ratio between the solid line and the dots in Figs. 24(a) and 24(b). The resulting $\varepsilon_\infty(X)$ is plotted in Fig. 26 [dots and crosses refer to Figs. 24(a) and 24(b), respectively]. According to Coulomb gas scaling (see Sec. IV.D.2), $\varepsilon_\infty(X)$ should be a "universal" (i.e., sample-independent) function. As can be seen in Fig. 26 this Coulomb gas feature of the data is rather well fulfilled.

An interesting question in this context is to what extent the functional form of the $\varepsilon_\infty(X)$ extracted from the nonlinear IV characteristics can be tied to the Coulomb gas model. The solid curve in Fig. 26 is a calculation based on the Coulomb gas renormalization equations [Eqs. (3.14); Minnhagen, 1986]. The starting point for this calculation is the Ginzburg-Landau Coulomb gas model (see Sec. IV.D and Table IV), supplemented with the value $\tilde{c}\simeq 13$ obtained from the "universal" resistance curve (Minnhagen and Weber, 1985) and the value $\varepsilon_c\simeq 1.65$ obtained from nonlinear IV characteristics (Minnhagen,

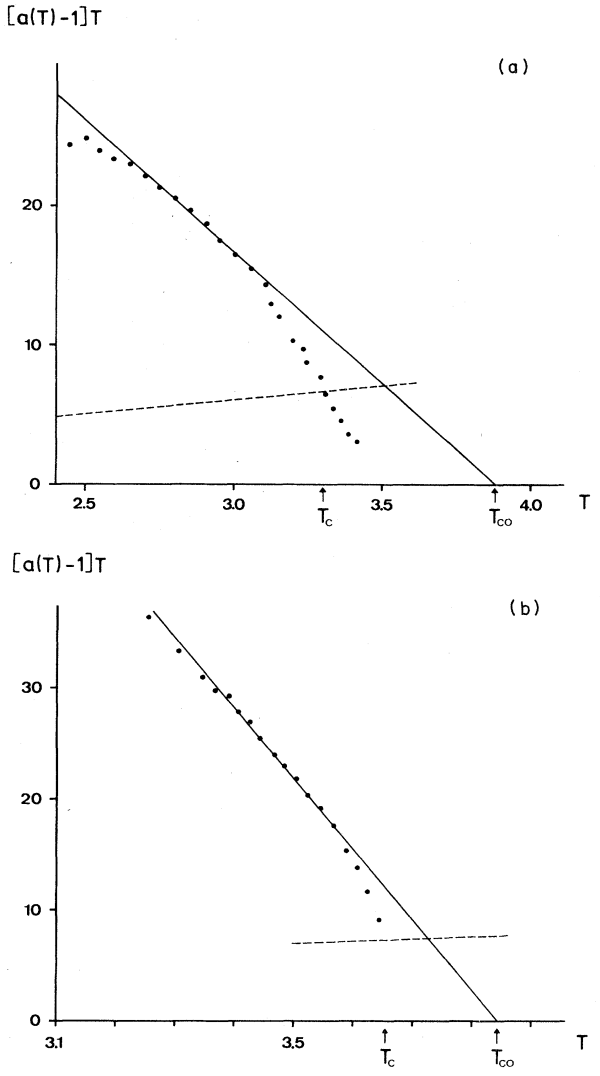

FIG. 24. The exponent $a(T)$ for the nonlinear IV characteristics. Measurements in (a) and (b) are from two Hg-Xe-alloy samples (represented by dots, data from Kadin *et al.*, 1983; for details, see Minnhagen, 1983b). The data are plotted as $[a(T)-1]T$ vs T. For lower temperatures a linear behavior $[a(T)-1]T\sim T_{c0}-T$ is expected [solid lines in (a) and (b)]. Crossing between solid line and T axis gives T_{c0}. The dashed lines are $[a(T)-1]T=2T$. Crossing between dashed line and measured values gives T_c. [The determined values for T_c and T_{c0} are $T_c=3.30$ K, $T_{c0}=3.88$ K for case (a); $T_c=3.64$ K, $T_{c0}=3.84$ K for case (b).]

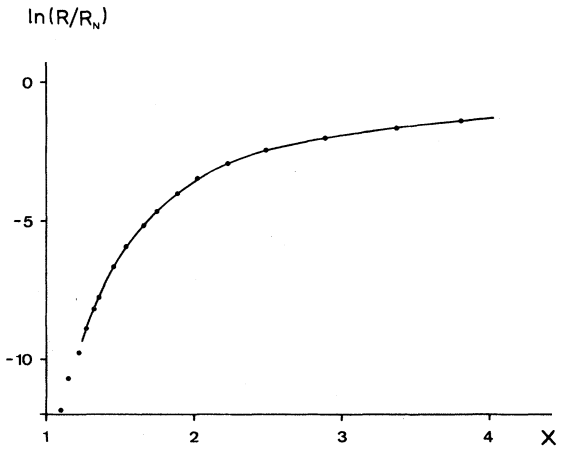

FIG. 25. Example of a consistency test. The resistance ratio for the same sample as in Fig. 24(a) is plotted against $X=T(T_{c0}-T_c)/[T_c(T_{c0}-T)]$, where T_{c0} and T_c are the same as in Fig. 24(a). The resistance data are represented by dots. The solid curve is the "universal" resistance curve (from Minnhagen, 1983b).

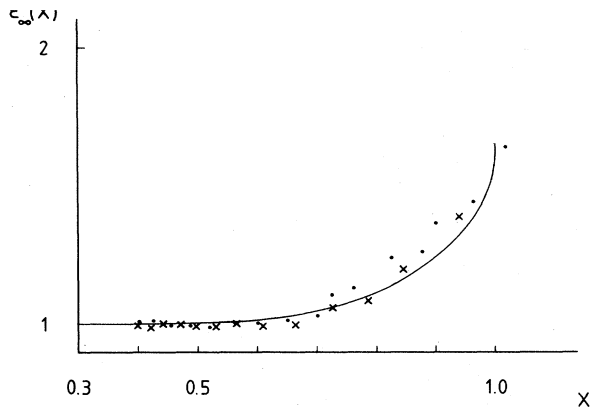

FIG. 26. The Coulomb gas dielectric constant $\varepsilon_\infty(X)$ extracted from the data given by Fig. 24. Dots and crosses refer to Figs. 24(a) and 24(b), respectively. The ε_∞ extracted from the two samples agree rather well. The solid curve is a calculation based on the Coulomb gas renormalization equations (see text).

1983b). The renormalization equations are integrated from $l_i[=\ln(r/2r_0)]=0$. The initial values are

$$z_i=\frac{(2r_0)^2}{\zeta}e^{-E_c/(4XT_c^{CG})}$$

and $T_i^{CG}=\varepsilon_i T_c^{CG}X$, with $E_c=0.390$, $\zeta=\tilde{c}r_0^2\simeq 13r_0^2$, and $T_c^{CG}=1/(4\varepsilon_c)\simeq 0.152$. The initial value $\varepsilon_i(T_c)\simeq 1.39$ follows from the renormalization equations and the condition $T_c^{CG}\simeq 0.152$. The rest of the $\varepsilon_i(T)$ values may be estimated by $\varepsilon_i=(1-2\pi\chi_i)^{-1}\simeq 1+2\pi\chi_i$, where χ_i, the polarizability due to neutral pairs with separation less than $2r_0$, is assumed to be proportional to the density of such pairs, which in turn is expected to be proportional to z_i^2 for small densities. This gives the estimate

$$\varepsilon_i(T)\simeq 1+[\varepsilon_i(T_c)-1][z_i(T)/z(T_c)^2]$$

used in the calculation. When comparing this calculated Coulomb gas $\varepsilon_\infty(X)$ with the experimentally determined $\varepsilon_\infty(X)$ it should be noted that the agreement at $X=1$ is guaranteed [because $\varepsilon_c\simeq 1.65$ used in the calculation was obtained from the experimental $\varepsilon_c(X)$]. Nevertheless, as is apparent from Fig. 26, the agreement between this Coulomb gas calculation and the measurements is rather suggestive.

In summary the situation for the connection between the nonlinear IV characteristics and the Coulomb gas model seems to be the following: The Coulomb gas prediction $a(X=1)=3$ at T_c is borne out (as is especially clearly demonstrated by Hebard and Fiory, 1983, and Fiory *et al.*, 1983). The Coulomb gas interpretation of the nonlinear IV characteristics is consistent with the Coulomb gas interpretation of the complex impedance close to T_c (Fiory *et al.*, 1983). However, below T_c inconsistencies in the Coulomb gas interpretation often appear, possibly connected with vortex pinning effects. For some samples, like the Hg-Xe alloy samples discussed above, the Coulomb gas interpretation appears to be consistent in a temperature region extending somewhat below T_c. An example of evidence in this direction is the agreement between the Coulomb gas interpretations of the nonlinear IV characteristics below T_c and the flux-flow resistance above T_c (Minnhagen, 1983b). Another is the apparent Coulomb gas scaling of the $\varepsilon_\infty(X)$ extracted from experiments (Minnhagen, 1983b).

3. Magnetic field scaling of the resistance

According to the Coulomb gas model, the resistance ratio R/R_N should (in the limit of a small external current j_{ex}) be a "universal" (i.e., sample-independent) function of the two variables X and B_{ex}/B_{sc}, where B_{ex} is an external perpendicular magnetic field and $B_{sc}(=\varphi_0/2\pi\xi^2)$ is the scaling magnetic field [see Sec. IV.D.2 in connection with Eq. (4.60)]. This Coulomb gas scaling appears to be fulfilled by experiments (Minnhagen, 1983a, 1984b; Lee, 1985).

Figure 27 gives a direct illustration of the Coulomb gas scaling (Lee, 1985). The resistance ratios R/R_N in the

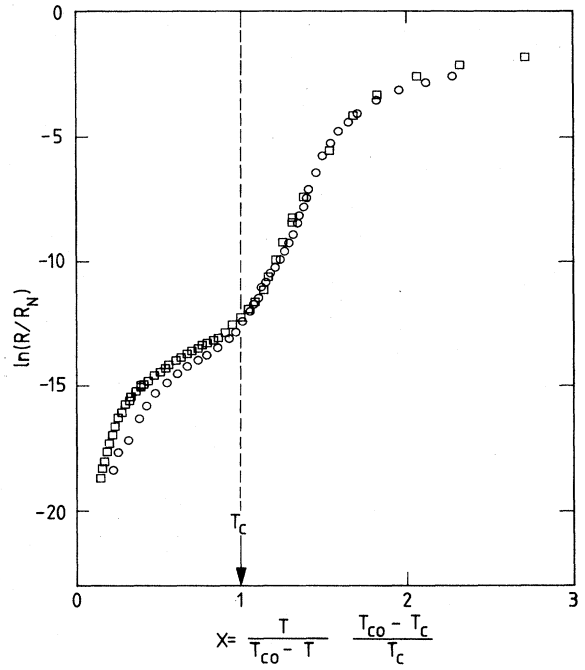

FIG. 27. Resistance ratio for two indium/indium oxide samples in a fixed perpendicular magnetic field $B_{ex}=50$ mG. The resistance is plotted against the scaling variable $X=T(T_{c0}-T_c)/[T_c(T_{c0}-T)]$. The data for the two samples are represented by squares and circles. T_{c0} and T_c are determined from the Aslamozov-Larkin formula and the nonlinear IV characteristics, respectively (for details see Lee, 1985: □, $T_c=3.122$ K, $T_{c0}=3.298$ K; ○, $T_c=2.727$ K, $T_{c0}=2.990$ K). The scaling fields $B_{sc}(X)$ are almost equal for the two samples. The agreement above T_c may be interpreted as a manifestation of Coulomb gas scaling (from Lee, 1985).

case of a fixed perpendicular magnetic field $B_{ex}=50$ mG for two indium/indium oxide samples are plotted against the scaling variable $X=T(T_c-T_{c0})/[T_c(T_{c0}-T)]$. The temperatures T_{c0} and T_c are in this case independently determined from the Aslamozov-Larkin formula (Aslamozov and Larkin, 1968) and the nonlinear IV characteristics (compare Sec. V.B.2), respectively. The scaling fields $B_{sc}(X)$ are almost equal for the two samples. Consequently the overlap of the data above T_c can be interpreted as a manifestation of Coulomb gas scaling (Lee, 1985). Below T_c this scaling property appears to break down, possibly due to vortex pinning effects (Lee, 1985).

Figure 28 gives another illustration (Minnhagen, 1984b). The solid curve in the figure represents a measurement taken at a fixed B_{ex} as a function of temperature (an aluminum/aluminum oxide sample measured by Masker *et al.*, 1969). In Fig. 28 the measured resistance ratio $R/R_N\,|_{B_{ex}}$ is plotted as a function of the corresponding resistance ratio for $B_{ex}=0$, i.e., $R/R_N\,|_{B_{ex}=0}$. The four dots represent four samples measured at fixed temperature as a function of B_{ex} (four indium/indium oxide samples measured by Fiory *et al.*, 1983). The dots give the measured values of these samples at the same value of the scaling variable B_{ex}/B_{sc} as the sample represented by the solid curve. Magnetic field scaling is obeyed provided the dots fall on the solid curve and is obviously well borne out by the example given in Fig. 28.

The Coulomb gas model leads to the following functional expression for the flux-flow resistance (Minnhagen, 1981a): The density of free Coulomb gas particles for a non-neutral Coulomb gas is given by Eq. (2.44). By use of the Coulomb gas analogy this can be translated [compare Sec. IV.D.2 in connection with Eq. (4.60)] to

$$\frac{B_{ex}}{B_{sc}}=\frac{R}{R_N}\left[1-\kappa^2\left[\frac{R}{R_N}\right]^{-2/\nu}\right]^{1/2}, \tag{5.6}$$

where κ and ν are functions of X and B_{ex}/B_{sc} and ν is related to the Coulomb gas dielectric constant by $\nu=4\tilde{\epsilon}T^{CG}/(4\tilde{\epsilon}T^{CG}-1)$. However, the functional forms of the functions $\kappa(X,B_{ex}/B_{sc})$ and $\nu(X,B_{ex}/B_{sc})$ are at present unknown.

Equation (5.6) predicts a crossover from a high-temperature region where the flux-flow resistance is dominated by thermally created free vortices to a low-temperature region where the flux-flow resistance is dominated by the vortices corresponding to the external field (compare Fig. 4, which illustrates this crossover in terms of the density of free Coulomb gas particles). The crossover behavior described by Eq. (5.6) is in qualitative agreement with measurements (Fiory *et al.*, 1983; Minnhagen, 1983a; Kihmi *et al.*, 1984; Lee, 1985). However, a more critical test of Eq. (5.6) by experiments has so far not been achieved; one difficulty is that the functional forms of κ and ν are at present unknown.

Thus, in summary, the magnetic field scaling predicted from the Coulomb gas analogy appears to be borne out by experiments. The functional form of the flux-flow resistance predicted from the Coulomb gas model is in qualitative agreement with experiments, but a critical test of this functional form by experiments is lacking at present.

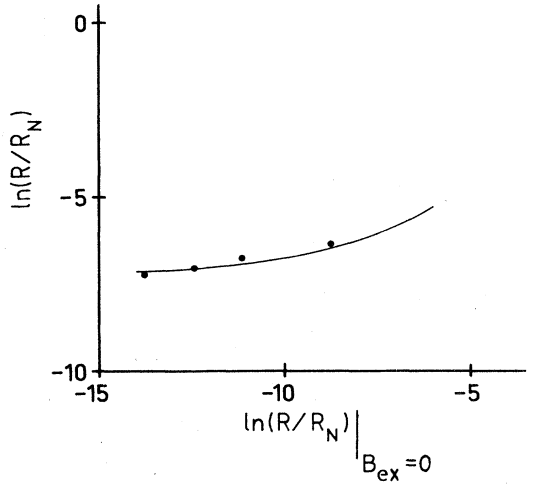

FIG. 28. Example of magnetic field scaling for resistance (from Minnhagen, 1984b). The solid curve is an aluminum/aluminum oxide sample measured for $B_{ex}=0.37$ G. The resistance ratio is plotted as a function of the corresponding $B_{ex}=0$ resistance ratio. The four dots represent measurements of four indium/indium oxide samples for the same values of the magnetic scaling variable B_{ex}/B_{sc} as the solid curve. According to Coulomb gas scaling, the dots should fall on the solid curve.

VI. DYNAMICAL ASPECTS

In the preceding section, experiments reflecting 'static" Coulomb gas properties were discussed. This section deals with modifications of the "static" Coulomb gas model aimed at a description of experimental results reflecting the dynamical aspects of vortices for superfluid-superconducting films.

The dynamics of the Coulomb gas is introduced by assuming that the Coulomb gas charges move according to a Langevin equation (Sec. VI.A). A central quantity in the dynamical description is the frequency-dependent complex dielectric function describing the response of the Coulomb gas to an external time-dependent electric field. A heuristic estimate of this complex dielectric function is given in Sec. VI.A.1. The imaginary part of the complex dielectric function describes the energy dissipation. This dissipation constitutes the most direct way in which the dynamics of the Coulomb gas is reflected in experiments on superfluid-superconducting films.

A description of vortex dynamics for superfluid films was developed by Ambegaokar *et al.* (1978), Ambegaokar and Teitel (1979), and Ambegaokar *et al.* (1980). It involves a combination of the length-dependent screening reasoning (see Sec. III.B), a Langevin equation describing the dynamics, and Kosterlitz renormalization-group equations [Eqs. (3.16)]. Some aspects of this description are

treated in Sec. VI.A.2 within the Coulomb gas formulation. The relation to vortex dynamics (Ambegaokar *et al.*, 1978, 1980) is described in Sec. VI.A.3.

Section VI.B illustrates the "dynamical" Coulomb gas aspects by two experimental examples. The first concerns the dissipation observed in the torsion pendulum experiment for helium films (Sec. VI.B.1). The "static" aspects of this experiment were described in Sec. V.A. The second concerns the observed scaling properties for the complex dielectric function extracted from a superconducting film (Sec. VI.B.2).

Section VI.C discusses the connection between Maxwell's equations in two dimensions and vortex dynamics (Ambegaokar *et al.*, 1980).

A. Coulomb gas with dynamics

Dynamics is introduced into the "static" Coulomb gas model described in Secs. II and III by postulating that a Coulomb gas charge moves according to the Langevin equation

$$\frac{d\mathbf{r}_j(t)}{dt}=\frac{\mathcal{D}}{T^{\mathrm{CG}}}\mathbf{F}_j^{\mathrm{eff}}(t)+\boldsymbol{\eta}(t)\ , \tag{6.1}$$

where $\mathbf{r}_j$ is the position vector of a Coulomb gas particle, $\mathbf{F}_j^{\mathrm{eff}}$ is the effective force acting on it, $\mathcal{D}$ is the diffusion constant for Coulomb gas charges [with dimension (length)^2/time], and $\boldsymbol{\eta}$ is a Gaussian noise obeying

$$\langle \eta^{\alpha}(t)\eta^{\beta}(t')\rangle = 2\mathcal{D}\delta_{\alpha\beta}\delta(t-t')\ , \tag{6.2}$$

where α and β denote the Cartesian components. This Langevin equation for the Coulomb gas mimics the vortex dynamics for superfluid films (Ambegaokar *et al.*, 1978, 1980; see Sec. VI.A.3). The aim is to describe the situation in which a small time-dependent external electric field $\mathbf{D}(t)$ is acting on the Coulomb gas. The time dependence of $\mathbf{D}(t)$ is taken to be an oscillation with frequency ω, i.e., $\mathbf{D}(t)=\mathbf{D}^0\cos(\omega t)$.

Linear-response theory can be used, since the aim is a description valid for small D. Consequently the effective mean electric field $\mathbf{E}^{\mathrm{eff}}(t)$ acting on the Coulomb gas charges can be expressed in terms of a complex-valued frequency-dependent dielectric function $\varepsilon(\omega)=\varepsilon_R(\omega)+i\varepsilon_I(\omega)$, i.e.,

$$\mathbf{E}^{\mathrm{eff}}(t)=\mathbf{D}^0\mathrm{Re}\left\{\frac{e^{-i\omega t}}{\varepsilon(\omega)}\right\}\ , \tag{6.3}$$

where Re{ } denotes the real part of { }.

The dielectric function $\varepsilon(\omega)$ will be a fundamental quantity in the description. It is related to the time-averaged superfluid density ρ_S by [compare Eq. (4.43)]

$$\frac{\rho_S}{\rho}=\mathrm{Re}\left\{\frac{1}{\varepsilon(\omega)}\right\}\ . \tag{6.4}$$

This relation can be inferred in the following way: First one observes [see Eq. (6.3)] that

$$\mathrm{Re}\left\{\frac{1}{\varepsilon(\omega)}\right\}=\int_0^{t_p}\frac{dt}{t_p}\frac{E^{\mathrm{eff}}(t)}{D(t)}\ , \tag{6.5}$$

where the time average is taken over a period $t_p=2\pi/\omega$. The force acting on a vortex for a neutral (charged) superfluid is the Magnus (Lorentz) force, which is proportional to the superfluid mass flow (see Sec. VI.A.3). Thus $\mathbf{D}(t)$ is proportional to an externally imposed mass current $\mathbf{g}_{\mathrm{ex}}(t)$ and $\mathbf{E}^{\mathrm{eff}}(t)$ to the resulting average of the total mass current $\mathbf{g}_{\mathrm{tot}}(t)$, so that

$$\frac{g_{\mathrm{tot}}(t)}{g_{\mathrm{ex}}(t)}=\frac{E^{\mathrm{eff}}(t)}{D(t)}\ . \tag{6.6}$$

As discussed in Sec. IV.A.1, the superfluid density can be defined by Eqs. (4.9), which in the static case leads to an expression in terms of mass-current correlation functions given by Eq. (4.10). This expression can alternatively be cast into the form (Minnhagen and Warren, 1981)

$$\frac{\rho_S}{\rho}=\lim_{g_{\mathrm{ex}}\to 0}\frac{g_{\mathrm{tot}}}{g_{\mathrm{ex}}}\ , \tag{6.7}$$

which means that the ratio between the renormalized and "bare" superfluid density is equal to the ratio of the total and imposed mass currents in the limit of a small imposed mass current. This definition of the superfluid density carries over to the time-dependent case, which means that the superfluid density in this case is time dependent and given by

$$\frac{\rho_S(t)}{\rho}=\frac{g_{\mathrm{tot}}(t)}{g_{\mathrm{ex}}(t)}\quad \text{for } g_{\mathrm{ex}}\to 0\ . \tag{6.8}$$

The time-averaged superfluid density ρ_S is just

$$\rho_S=\int_0^{t_p}\frac{dt}{t_p}\rho_S(t)\ , \tag{6.9}$$

and Eq. (6.4) now follows from Eqs. (6.5)—(6.9).

In Sec. VI.A.1 we obtain an expression for $\varepsilon(\omega)$ by a heuristic argument and discuss the qualitative features of $\varepsilon(\omega)$. Section VI.A.2 describes some aspects of the conjunction between the length-dependent screening reasoning and dynamics, while Sec. VI.A.3 makes contact with vortex dynamics.

1. Qualitative features of the dielectric function

The Langevin equation for the Coulomb gas charges [Eq. (6.1)] introduces an additional length into the description, namely, a characteristic diffusion length $\sqrt{\mathcal{D}/\omega}$ (Ambegaokar *et al.*, 1978). One may heuristically obtain approximations for the dielectric function $\varepsilon(\omega)$ describing the dynamics directly from a static description by focusing on the role played by this diffusion length (Ambegaokar *et al.*, 1978).

One version of this type of reasoning is the following: Let us start from the static result in the form [compare Eqs. (3.4) and (4.43)]

$$\frac{\rho_S}{\rho}=1+\frac{\pi^2}{T^{\rm CG}}\int_0^\infty dr\, r^3\langle \Delta n(r)\Delta n(0)\rangle \ . \tag{6.10}$$

As discussed in Sec. III.B, the right-hand side can be intuitively interpreted as describing the polarization caused by neutral pairs, where r is the separation between the particles in a pair. Now, let us try to introduce dynamics into this intuitive picture. Obviously the particles in a pair with separation r have to move a distance $2\pi r$ during a period if the pair is going to maintain a fixed orientation with respect to the electric field. This means that only pairs with separation $r \ll \sqrt{D/\omega}$ will have time to relax in the electric field. The contribution from such pairs to the polarization is just the same as for the static case. On the other hand, pairs with separation $r \gg \sqrt{D/\omega}$ will not have time to relax in the electric field and hence will not, when time-averaged over a period, contribute to the polarization. This suggests that the time-averaged superfluid density for the dynamical case is approximately given by

$$\frac{\rho_S}{\rho}=1+\frac{\pi^2}{T^{\rm CG}}\int_0^{\sqrt{D/\omega}} dr\, r^3\langle \Delta n(r)\Delta n(0)\rangle \ , \tag{6.11}$$

where $\langle\ \rangle$ denotes the thermal average for the static Coulomb gas. Expressed in terms of the corresponding Fourier transform [Eq. (4.42)], this in turn suggests that ρ_S for the dynamical case is approximately given by

$$\frac{\rho_S}{\rho}=\left[1-\frac{\hat{V}(k)}{T^{\rm CG}}\langle \Delta\hat{n}(\mathbf{k})\Delta\hat{n}(-\mathbf{k})\rangle\right]\Bigg|_{k=1/\lambda_\omega} , \tag{6.12}$$

where λ_ω is a length of the order of $\sqrt{D/\omega}$. This may alternatively, by use of Eq. (4.45), be expressed as

$$\frac{\rho_S}{\rho}=\frac{1}{\tilde{\varepsilon}}\frac{1+(\lambda_\omega/\lambda_c)^2}{1+(\lambda_\omega/\lambda)^2} \ , \tag{6.13}$$

which also covers the generalization to the case of a nonneutral Coulomb gas (Minnhagen, 1981a). The length λ_ω will be referred to as the dynamical length; it may be interpreted as a measure of the length scale over which the time average of the charge-density correlations in the Coulomb gas is destroyed due to the dynamics. The length λ_c is the cutoff length of the "bare" particle interaction [compare Eqs. (2.3) and (2.5)] and λ is the effective screening length [compare Eqs. (2.21), (2.22), and (2.25)].

The superfluid density is related to the dielectric function by Eq. (6.4), so that Eq. (6.13) is really an expression for $\mathrm{Re}\{1/\varepsilon(\omega)\}$. Consequently the imaginary part of the same response function, i.e., $\mathrm{Im}\{1/\varepsilon(\omega)\}$, can be obtained by the Kramers-Kronig relation

$$\mathrm{Im}\left\{\frac{1}{\varepsilon(\omega)}\right\}=\frac{1}{\pi}P\int_0^\infty d\omega'\left[\mathrm{Re}\left\{\frac{1}{\varepsilon(\omega')}\right\}-1\right]\times\left[\frac{2\omega}{\omega^2-\omega'^2}\right] , \tag{6.14}$$

where P denotes the principal part. Since λ_ω^2 is proportional to ω, an explicit expression for $\mathrm{Im}\{1/\varepsilon(\omega)\}$ is readily obtained from Eq. (6.14) by using $\mathrm{Re}\{1/\varepsilon(\omega)\}$ given by Eq. (6.13) as input. The result is

$$\mathrm{Re}\left\{\frac{1}{\varepsilon(\omega)}\right\}=\frac{1}{\tilde{\varepsilon}}\frac{1+(\lambda_\omega/\lambda_c)^2}{1+(\lambda_\omega/\lambda)^2} \ , \tag{6.15a}$$

$$\mathrm{Im}\left\{\frac{1}{\varepsilon(\omega)}\right\}=\frac{4}{\pi\tilde{\varepsilon}}[(\lambda_\omega/\lambda)^2-(\lambda_\omega/\lambda_c)^2]\times\ln(\lambda_\omega/\lambda)[1-(\lambda_\omega/\lambda)^4]^{-1} \ . \tag{6.15b}$$

An appealing feature of this expression is that it describes the qualitative physics in terms of the interplay between three length scales, i.e., the screening length λ, the dynamical length λ_ω, and the range of the "bare" particle interaction λ_c.

The dominant qualitative physics reflected in the quantity $1/\varepsilon(\omega)$ is determined by the rapid decrease of the screening length λ with increasing temperature due to charge unbinding (see Sec. II.B). Consequently, on a qualitative level, the temperature variation of the dielectric constant $\tilde{\varepsilon}$, which describes the polarization due to bound pairs, can be neglected and $\tilde{\varepsilon}$ can be replaced by its "bare" value $\tilde{\varepsilon}=1$ (see Sec. II.B). Equations (6.15) then reduce to

$$\mathrm{Re}\left\{\frac{1}{\varepsilon(\omega)}\right\}=\frac{1+\bar{C}}{1+Y} \ , \tag{6.16a}$$

$$\mathrm{Im}\left\{\frac{1}{\varepsilon(\omega)}\right\}=\frac{4(Y-\bar{C})\ln(Y)}{\pi(1-Y^2)} \ , \tag{6.16b}$$

where $Y=(\lambda_\omega/\lambda)^2$ and $\bar{C}=(\lambda_\omega/\lambda_c)^2$.

A qualitative estimate of the actual temperature dependence can be obtained by approximating the rapid temperature variation of λ with the Coulomb gas estimate [see Eqs. (2.29) and (2.30)],

$$(\lambda/r_0)^{-2}=\{\mathrm{const}/[1-(\lambda/\lambda_c)^2]\}^{X/(X-1)} \ , \tag{6.17}$$

where $X=T^{\rm CG}/T_c^{\rm CG}$ is the Coulomb gas temperature variable [compare Eq. (4.59a)]. The dynamical length λ_ω can for this purpose be taken to be a temperature-independent constant.

Let us first consider the case in which the "bare" particle interaction is logarithmic to arbitrarily large distances, i.e., $\lambda_c=\infty$ [compare Eq. (2.5)]. This case is described by Eq. (6.16) with $\bar{C}=0$. The basic features given by Eqs. (6.16) for this case are that $\mathrm{Re}\{1/\varepsilon(\omega)\}$ drops from its T_c value to zero with increasing Y, whereas $-\mathrm{Im}\{1/\varepsilon(\omega)\}$ is a positive function that goes to zero for $Y\to 0^\infty$ and has a maximum for $Y=1$, i.e., has a maximum for the temperature for which the screening length λ is equal to the dynamical length λ_ω. [For lower (higher) temperatures the screening length λ rapidly becomes much larger (smaller) than λ_ω; compare Eq. (6.17)]. One notes that within the approximation given by Eqs. (6.16) the position of this maximum is precisely where $\mathrm{Re}\{1/\varepsilon(\omega)\}$ has

dropped to half its T_c value.

Figure 29 shows the qualitative temperature dependence of $1/\varepsilon(\omega)$ obtained from Eqs. (6.16) and (6.17) for three different values of λ_ω. The qualitative features illustrated in the figure are the following: The apparent transition region for $\mathrm{Re}\{1/\varepsilon(\omega)\}$ moves to higher temperatures and becomes broader with decreasing λ_ω. Likewise, the peak position of $-\mathrm{Im}\{1/\varepsilon(\omega)\}$ moves to higher temperatures and the peak becomes broader with decreasing λ_ω. However, within the approximation given, the peak height is independent of λ_ω and the peak position is always where $\mathrm{Re}\{1/\varepsilon(\omega)\}$ has half its T_c value, which corresponds to the temperature when the screening length λ is equal to the dynamical length λ_ω. The width of the peak shrinks to zero as λ_ω goes to infinity.

As will be shown in Sec. VI.B.1, $-\mathrm{Im}\{1/\varepsilon(\omega)\}$ describes the dissipation observed in the experiments, so that the qualitative features described above are rather directly open to experimental verifications.

One may also note that on the most crude level the essence of the description is that the dissipation is largest when the screening length due to free vortices is of the same order as the dynamical length λ_ω.

Next we consider the qualitative effects of the length scale λ_c on $1/\varepsilon(\omega)$. These are contained in Eq. (6.16) with $C\neq 0$ and can be extracted by using the same approximations as for the $\lambda_c=\infty$ case. Figure 30 shows how the $\lambda_\omega=\lambda_{\omega_2}$ case in Fig. 29 is modified by finite λ_c values. As is illustrated in Fig. 30, one may distinguish between three cases characterized by the relative sizes of the two length scales λ_c and λ_ω. The basic features are the following: The transition region for $\mathrm{Re}\{1/\varepsilon(\omega)\}$ becomes broader with decreasing λ_c. This is to a large extent just a reflection of the fact that the transition region for the screening length λ becomes larger (see Sec. II.B in connection with Fig. 2). The peak height of $-\mathrm{Im}\{1/\varepsilon(\omega)\}$ decreases with decreasing λ_c, and $\mathrm{Im}\{1/\varepsilon(\omega)\}$ rapidly vanishes altogether when λ_c becomes smaller than λ_ω. For $\lambda_c>\lambda_\omega$ the main effect of a finite λ_c on $-\mathrm{Im}\{1/\varepsilon(\omega)\}$ is to open up the peak towards lower temperatures.

The above description is based on a minimum of as-

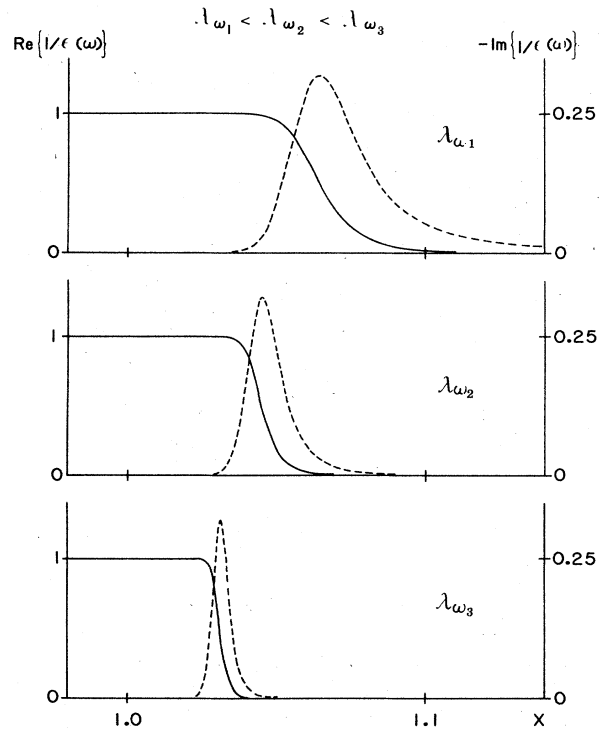

FIG. 29. The temperature dependence of the complex dielectric constant $\varepsilon(\omega)$ for the two-dimensional Coulomb gas, as given by the approximate equations (6.16) and (6.17). The figure illustrates the $\lambda_c=\infty$ case for three different values of λ_ω: $\lambda_{\omega_1}<\lambda_{\omega_2}<\lambda_{\omega_3}$ (the relative magnitudes chosen in the figure are $\lambda^2_{\omega_1}:\lambda^2_{\omega_2}:\lambda^2_{\omega_3}=1:10^2:10^5$). The real parts (solid curves) and the imaginary parts (dashed curves) of $1/\varepsilon(\omega)$ are plotted as functions of the Coulomb gas temperature variable $X=T^{\mathrm{CG}}/T_c^{\mathrm{CG}}$. The main features are described in the text. One may note that $\lambda_\omega/\lambda=[\mathrm{Re}\{1/\varepsilon(\omega)\}^{-1}-1]^{1/2}$ [compare Eq. (6.16a)].

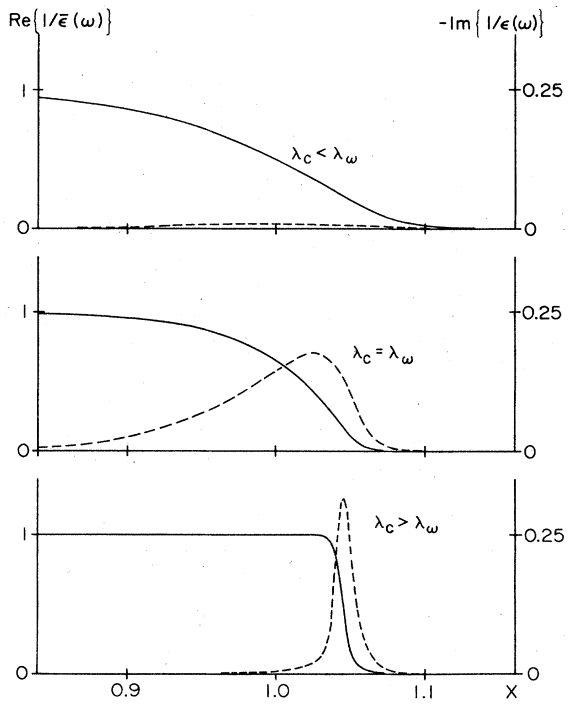

FIG. 30. The temperature dependence of the complex dielectric constant $\varepsilon(\omega)$ for the two-dimensional Coulomb gas as given by the approximate equations (6.16) and (6.17). The figure illustrates the $\lambda_c\neq\infty$ case for three different values of λ_c, i.e., $\lambda_c/\lambda_c<1$, $\lambda_c/\lambda_\omega=1$, and $\lambda_c/\lambda_\omega>1$ (the three λ_c/λ_ω ratios chosen in the figure are $\lambda_c^2/\lambda_\omega^2=10^{-2}$, 1, and 10^3, respectively. The other parameters are the same as for the λ_{ω_2} case of Fig. 30). The real parts (solid curves) and the imaginary parts (dashed curves) of $1/\varepsilon(\omega)$ are plotted as functions of the Coulomb gas temperature variable $X=T^{\mathrm{CG}}/T_c^{\mathrm{CG}}$. The main features are described in the text.

sumptions and focuses on the interplay between the three large length scales involved, i.e., λ, λ_ω, and λ_c. Consequently, the features obtained may be expected to be the true qualitative ones, which should be qualitatively independent of more detailed model assumptions.

In Sec. IV.A.2 a more "canonical" version of the dynamics is outlined. This description is due to Ambegaokar *et al.* (1978, 1980) and Ambegaokar and Teitel (1979) and covers the $\lambda_c=\infty$ case for a neutral Coulomb gas. It involves more explicit dynamical assumptions, but as expected the qualitative features obtained agree well with the description given above.

Another version of the dynamical theory has been given by Shenoy (1985a, 1985b). This description starts from a microscopic model of a square array of coupled Josephson junctions. Again, on a qualitative level, the result obtained agrees with the $\lambda_c=\infty$ case for the neutral Coulomb gas given above.

The $\lambda_c\neq\infty$ case for a neutral Coulomb gas has been considered by Kotsubo and Williams (1984, 1986) and Wang and Lu (1986). These authors study the very special model of a Coulomb gas confined on a spherical surface with radius $\mathscr{R}$. The radius $\mathscr{R}$ plays the role of the cutoff λ_c for the "bare" particle interaction. With this obvious identification the qualitative features obtained agree well with the description given in the present section.

2. Length-dependent screening and dynamics

This section describes a reasoning that more directly combines the length-dependent screening idea (Sec. III.B) and the Langevin equation (6.1), the object being to obtain an approximation for the dielectric constant $\varepsilon(\omega)$. The reasoning is due to Ambegaokar *et al.* (1978, 1980) and Ambegaokar and Teitel (1979). It concerns the case of a neutral Coulomb gas with an infinite range for the "bare" particle interaction, i.e., $\lambda_c=\infty$.

The reasoning is based on the intuitive picture that the Coulomb gas consists of a mixture of free charges and neutral bound pairs of charges. In accordance with this picture, it is assumed that the motion of a free charge is only weakly correlated with the motion of other individual charges. In contrast, the motion of one of the charges of a neutral bound pair is strongly correlated with the motion of the other member of the pair. The free charges and the bound pairs will then give two additive contributions to the complex electric susceptibility, $\chi(\omega)$, so that [compare Eqs. (3.6)—(3.8)]

$$\varepsilon(\omega)=1+2\pi\chi(\omega)=1+2\pi\chi_F(\omega)+2\pi\chi_B(\omega)\ , \quad (6.18)$$

where the subscript F (B) denotes free (bound) charges.

The effective force $\mathbf{F}_j^{\rm eff}$ acting on a *free* particle is approximately given by the mean electric field $\mathbf{E}^{\rm eff}$ because the free-particle motion is only weakly correlated to other individual particles, i.e.,

$$\mathbf{F}_j^{\rm eff}=s_j\mathbf{E}^{\rm eff}\ , \quad (6.19)$$

where s_j is the charge of the free particle. From the Langevin equation (6.1) and the linear-response result [Eq. (6.3)] one then obtains

$$\sum_j s_j\langle\mathbf{r}_j(t)\rangle=-\frac{\mathscr{D}n_F D^0}{\omega T^{\rm CG}}{\rm Im}\left\{\frac{e^{-i\omega t}}{\varepsilon(\omega)}\right\}\ , \quad (6.20)$$

where n_F is the density of free charges and $\langle\ \rangle$ denotes the average. One can always fictitiously divide the free charges into dipole pairs. The sum of the dipole vectors in such a division is given by

$$\sum_{\substack{\rm dip\\ \langle ij\rangle}}(\mathbf{r}_i-\mathbf{r}_j)=\sum_j s_j\mathbf{r}_j\ . \quad (6.21)$$

The polarization is just the sum over the dipole vectors, which within a linear-response description is given by Eq. (6.20). This result can be expressed in terms of an electric susceptibility $\chi_0(\omega)$, which corresponds to the "independent dipole approximation" discussed in Sec. III.B, i.e.,

$$2\pi\chi_0(\omega)=1-1/\varepsilon(\omega)\ , \quad (6.22)$$

where

$$\chi_0(\omega)\mathbf{D}(\omega)=\frac{1}{\Omega}\sum_j s_j\langle\mathbf{r}_j(\omega)\rangle\ \text{ for } D(\omega)\to 0\ . \quad (6.23)$$

Here $\langle\mathbf{r}_j(\omega)\rangle$ is given by

$$\langle\mathbf{r}_j(\omega)\rangle=\int_0^{t_p}\frac{dt}{t_p}e^{i\omega t}\langle\mathbf{r}_j(t)\rangle \quad (6.24)$$

and similarly for $\mathbf{D}(t)$, so that $\mathbf{D}(\omega)=\mathbf{D}^0$ in the present case. Combining Eqs. (6.20)—(6.23) gives

$$\varepsilon(\omega)=1+2\pi i\sigma/\omega\ , \quad (6.25a)$$

where

$$\sigma=\frac{n_F\mathscr{D}}{T^{\rm CG}} \quad (6.25b)$$

and σ can be thought of as a conductivity. Thus, in summary, the reasoning for the free charges leads to a contribution to the electric susceptibility given by (Ambegaokar *et al.*, 1978)

$$\chi_F(\omega)=\frac{in_F\mathscr{D}}{\omega T^{\rm CG}}\ . \quad (6.26)$$

We now turn to the contribution from the bound neutral pairs. The electric susceptibility $\chi_0(\omega)$ can be expressed as [compare Eqs. (6.22)—(6.24)]

$$\chi_0(\omega)=\frac{1}{\Omega}\sum_j s_j\frac{\partial\langle\mathbf{r}_j(\omega)\rangle}{\partial\mathbf{D}(\omega)}\bigg|_{D=0}=\frac{1}{\Omega\varepsilon(\omega)}\sum_j s_j\frac{\partial\langle\mathbf{r}_j(\omega)\rangle}{\partial\mathbf{E}^{\rm eff}(\omega)}\bigg|_{E^{\rm eff}=0} \quad (6.27)$$

or [compare Eqs. (6.18) and (6.22)]

$$\chi_B(\omega)=\frac{1}{\Omega}\sum_j s_j\frac{\partial\langle\mathbf{r}_j(\omega)\rangle}{\partial\mathbf{E}^{\rm eff}(\omega)}\bigg|_{E^{\rm eff}=0}\ . \quad (6.28)$$

In the static case this reduces to

$$\chi_0=\frac{1}{\Omega}\sum_j s_j\frac{\partial\langle \mathbf{r}_j\rangle}{\partial \mathbf{D}}\bigg|_{D=0} \tag{6.29}$$

and [compare Eq. (6.21)]

$$\chi_B=\frac{1}{\Omega}\sum_j s_j\frac{\partial\langle \mathbf{r}_j\rangle}{\partial \mathbf{E}^{\rm eff}}\bigg|_{E^{\rm eff}=0}=\frac{1}{\Omega}\sum_{\substack{\rm dip\\ \langle ij\rangle}}\frac{\partial\langle \mathbf{r}_{ij}\rangle}{\partial \mathbf{E}^{\rm eff}}\bigg|_{E^{\rm eff}=0}. \tag{6.30}$$

Let us introduce the length-dependent screening reasoning (compare Sec. III.B) through a *static* length-dependent electric susceptibility $\chi_B(r)$ defined by [compare Eq. (6.30)]

$$\chi_B(r)=\int_0^r dr'\frac{\partial}{\partial r'}\chi_B(r')=\int d^2r'\theta(r-r')\frac{1}{\Omega}\times\sum_{\substack{\rm dip\\ \langle ij\rangle}}\frac{\partial\langle \mathbf{r}_{ij}\rangle}{\partial \mathbf{E}^{\rm eff}}\bigg|_{E^{\rm eff}=0}\delta(\mathbf{r}'-\mathbf{r}_{ij})\ . \tag{6.31}$$

The intuitive picture is that $(\partial/\partial r)\chi_B(r)$ gives the contribution to the electric susceptibility from all dipole pairs with separation r between the particles within a pair. If there are free charges present, then the particle interaction is screened out, so that the largest dipole pairs in the intuitive picture would have a separation of the order of the screening length λ. Consequently the contribution from the bound pairs to the electric susceptibility can, in the static case, using this intuitive picture, be estimated by $\chi_B(\lambda)$.

Next we consider how this picture is modified by the dynamics. The contribution to $\chi_B(\omega)$ can again be considered to be built up by contributions from the individual pairs (as is apparent by comparing Eqs. (6.28) and (6.30)]. Consequently the length-dependent electric susceptibility can be generalized to the dynamical situation by [compare Eq. (6.31)]

$$\chi_B(r,\omega)=\int d^2r'\theta(r-r')\frac{1}{\Omega}\times\sum_{\substack{\rm dip\\ \langle ij\rangle}}\frac{\partial\langle \mathbf{r}_{ij}(\omega)\rangle}{\partial \mathbf{E}^{\rm eff}(\omega)}\bigg|_{E^{\rm eff}=0}\delta(\mathbf{r}'-\mathbf{r}_{ij})\ , \tag{6.32a}$$

which we may rewrite as

$$\chi_B(r,\omega)=\int_0^r dr'\tilde{g}(r',\omega)\frac{\partial}{\partial r'}\chi_B(r')\ , \tag{6.32b}$$

where the function $\tilde{g}(r',\omega)$ is introduced to express the difference between the static and dynamical cases. The function $\tilde{g}(r',\omega)$ can be interpreted as expressing the modification from the static to the dynamical situation of the contribution to the electric susceptibility from pairs with intraparticle separation r'.

An argument by Ambegaokar and Teitel (1979) based on the Langevin equation [Eq. (6.1)] shows that the function $g(r,\omega)$ can be approximated by

$$\tilde{g}(r,\omega)=\frac{\lambda_\omega^2}{\lambda_\omega^2-ir^2}\ , \tag{6.33}$$

where λ_ω is a dynamical length scale given by (compare Sec. VI.A.1)

$$\lambda_\omega=\sqrt{14\mathcal{D}/\omega}\ . \tag{6.34}$$

The main steps in this argument are the following.

Consider the expression for the static charge-density correlation function given by Eq. (3.3a). For $r<\lambda$ this reduces to

$$\langle\Delta n(r)\Delta n(0)\rangle=-\frac{2z^2}{\zeta^2}\exp\{[V_L(r)-V_L(0)]/T^{\rm CG}\} \tag{6.35}$$

because $\{1-\exp[-2V_L(r)/T^{\rm CG}]\}$ is approximately equal to 1 for $r<\lambda$. $V_L(0)-V_L(r)$ is the work needed to separate two charges of opposite sign by a distance r, and $(d/d\mathbf{r})V_L(r)$ is the attractive effective force acting between them. An external field contributes an additional amount of work $-\mathbf{E}^{\rm eff}\cdot\mathbf{r}$, where $\mathbf{r}$ is the dipole vector associated with the pair. This corresponds to an additional repulsive force $\mathbf{E}^{\rm eff}$ between the particles [compare Eq. (2.48)]. The Langevin equation for a dipole vector is obtained by combining the Langevin equations for the two particles in the pair. This gives

$$\frac{d\mathbf{r}(t)}{dt}=\frac{2\mathcal{D}}{T^{\rm CG}}\mathbf{F}_{\rm dip}^{\rm eff}(t)+\boldsymbol{\eta}\ , \tag{6.36a}$$

which is the Langevin equation for a neutral pair with the dipole vector $\mathbf{r}$ [compare Eq. (6.1)],

$$\langle\eta^\alpha(t)\eta^\beta(t')\rangle=4\mathcal{D}\delta_{\alpha\beta}(t-t') \tag{6.36b}$$

describes the corresponding Gaussian noise η [compare Eq. (6.2)], and

$$\mathbf{F}_{\rm dip}^{\rm eff}(t)=-\frac{d}{d\mathbf{r}}\mathcal{V}(\mathbf{r},t) \tag{6.36c}$$

is the effective force acting between the two particles in a neutral pair, expressed in terms of a potential $\mathcal{V}(\mathbf{r},t)$. Within a linear-response description the potential $\mathcal{V}(\mathbf{r},t)$ is conveniently split into a zero-order and a first-order contribution in the external electric field, i.e.,

$$\mathcal{V}(\mathbf{r},t)=\mathcal{V}_0(\mathbf{r})+\mathcal{V}_1(\mathbf{r},t)\ . \tag{6.36d}$$

The zero-order contribution $\mathcal{V}_0(\mathbf{r})$ is just the static result

$$\mathcal{V}_0(\mathbf{r})=V_L(0)-V_L(\mathbf{r})\ , \tag{6.36e}$$

whereas the first-order contribution $\mathcal{V}_1(\mathbf{r},t)$ is given by

$$\mathcal{V}_1(\mathbf{r},t)=-\mathbf{E}^{\rm eff}(t)\cdot\mathbf{r}\ . \tag{6.36f}$$

As discussed in Sec. III.B, the charge-density correlation function can be interpreted in terms of the density of dipole pairs with separation $\mathbf{r}$. This dipole density will be denoted by $\Gamma(\mathbf{r},t)$. In the static case Γ is given by [com-

pare Eq. (3.5)]

$$\Gamma_0(\mathbf{r})=\mathrm{dip}(r)=-\Omega\langle\Delta n(\mathbf{r})\Delta n(0)\rangle/2 \sim e^{-\mathcal{V}(\mathbf{r})/T^{\mathrm{CG}}} \tag{6.37a}$$

and the generalization to the case of a small time-dependent external electric field is given by

$$\Gamma(\mathbf{r},t)=-\Omega\langle\Delta n(\mathbf{r},t)\Delta n(0,0)\rangle/2 =\Gamma_0(\mathbf{r})+\Gamma_1(\mathbf{r},t)\ , \tag{6.37b}$$

where Γ_1 is linear in the external field.

A Fokker-Planck equation for the dipole pair distribution $\Gamma(\mathbf{r},t)$ corresponding to the Langevin equation (6.36) is given by (Ambegaokar and Teitel, 1979; Teitel, 1981)

$$\frac{\partial}{\partial t}\Gamma(\mathbf{r},t)=\frac{2\mathcal{D}}{T^{\mathrm{CG}}}\frac{\partial}{\partial\mathbf{r}}\cdot\left[\Gamma(\mathbf{r},t)\frac{\partial}{\partial\mathbf{r}}\mathcal{V}(\mathbf{r},t)\right] +2\mathcal{D}\frac{\partial^2}{\partial\mathbf{r}^2}\Gamma(\mathbf{r},t)\ . \tag{6.38}$$

To linear order in the external field this becomes [compare Eqs. (6.36d)—(6.36f) and (6.37)]

$$-i\omega\Gamma_1(\mathbf{r},\omega)=\frac{2\mathcal{D}}{T^{\mathrm{CG}}}\frac{\partial}{\partial\mathbf{r}}\cdot\left[\Gamma_0(\mathbf{r})\frac{\partial}{\partial\mathbf{r}}\mathcal{V}_1(\mathbf{r},\omega) +\Gamma_1(\mathbf{r},\omega)\frac{\partial}{\partial\mathbf{r}}\mathcal{V}_0(\mathbf{r})\right] +2\mathcal{D}\frac{\partial^2}{\partial\mathbf{r}^2}\Gamma_1(\mathbf{r},\omega)\ , \tag{6.39a}$$

where [compare Eq. (6.24)]

$$\Gamma_1(\mathbf{r},\omega)=\int_0^{t_p}\frac{dt}{t_p}\Gamma_1(\mathbf{r},t)e^{i\omega t} \tag{6.39b}$$

and

$$\mathcal{V}(\mathbf{r},\omega)=\int_0^{t_p}\frac{dt}{t_p}\Gamma_1(\mathbf{r},t)e^{i\omega t} =-\mathbf{E}^{\mathrm{eff}}(\omega)\cdot\mathbf{r}\ . \tag{6.39c}$$

Since $\Gamma(\mathbf{r},t)$ is the distribution of dipole pairs, one has the obvious relation

$$\mathbf{r}\Gamma(\mathbf{r},t)=\frac{1}{\Omega}\sum_{\substack{\mathrm{dip}\\\langle ij\rangle}}\langle\mathbf{r}_{ij}(t)\rangle\delta(\mathbf{r}-\mathbf{r}_{ij}(t))\ , \tag{6.40}$$

and it follows that [compare Eq. (6.32a)]

$$\frac{\partial}{\partial r}\chi_B(r,\omega)=\int d^2r'\delta(r-r')\left.\frac{\partial[\mathbf{r}'\Gamma_1(\mathbf{r}',\omega)]}{\partial\mathbf{E}^{\mathrm{eff}}(\omega)}\right|_{E^{\mathrm{eff}}=0}\ . \tag{6.41}$$

Introducing the ansatz

$$\Gamma_1(\mathbf{r},\omega)=\mathbf{E}^{\mathrm{eff}}(\omega)\mathbf{r}\Gamma_0(r)\tilde{g}(r,\omega)/T^{\mathrm{CG}} \tag{6.42}$$

into Eq. (6.41) gives [compare Eqs. (3.5), (3.6), and (6.37a)]

$$\frac{\partial}{\partial r}\chi_B(r,\omega)=\tilde{g}(r,\omega)\frac{\partial}{\partial r}\chi_B(r)\ , \tag{6.43}$$

so that the $\tilde{g}(r,\omega)$ introduced by the ansatz is the same as was defined in Eq. (6.32). Consequently, by introducing the ansatz (6.42) into the Fokker-Planck equation (6.39), one obtains a differential equation for the function $\tilde{g}(r,\omega)$ defined by Eq. (6.32b). This differential equation is given by (Ambegaokar and Teitel, 1979)

$$r^2\frac{\partial^2}{\partial r^2}\tilde{g}+r\left[3-\frac{r}{T^{\mathrm{CG}}}\frac{\partial}{\partial r}\mathcal{V}_0\right]\frac{\partial}{\partial r}\tilde{g} -\left[\frac{r}{T^{\mathrm{CG}}}\frac{\partial}{\partial r}\mathcal{V}_0-\frac{i\omega r^2}{2\mathcal{D}}\right]\tilde{g}+\frac{r}{T^{\mathrm{CG}}}\frac{\partial}{\partial r}\mathcal{V}_0=0\ . \tag{6.44}$$

Close to the Kosterlitz-Thouless temperature T_c^{CG} this equation can be further simplified by using the approximation [compare Eqs. (6.36e), (3.10), and (3.21)]

$$\frac{r}{T^{\mathrm{CG}}}\frac{\partial}{\partial r}\mathcal{V}_0\simeq\lim_{r\to\infty}\left[\frac{r}{T^{\mathrm{CG}}}\frac{\partial}{\partial r}\mathcal{V}_0\right]\simeq\frac{1}{\varepsilon_\infty(T_c^{\mathrm{CG}})T_c^{\mathrm{CG}}}=4 \tag{6.45}$$

(provided T_c^{CG} is not too small; see Sec. III.D), and with this simplification Eq. (6.45) can be solved subject to the appropriate boundary conditions, $\tilde{g}(r,\omega=0)=1$ and $\tilde{g}(r,\omega=\infty)=0$. The solution is to good approximation given by Eqs. (6.33) and (6.34) (Ambegaokar and Teitel, 1979).

The general form of $\tilde{g}(r,\omega)$ given by Eq. (6.33) was suggested by Ambegaokar *et al.* (1978) on intuitive grounds. The intuitive physics implied by Eqs. (6.33) and (6.34) is that only dipole pairs with a pair separation less than λ_ω contribute to the electric susceptibility, because only such pairs have time to relax in the time-dependent electric field. This interpretation is apparent from the approximate equality (Ambegaokar *et al.*, 1980)

$$\mathrm{Re}\{\tilde{g}(r,\omega)\}=\frac{\lambda_\omega^4}{\lambda_\omega^4+r^4}\simeq\theta(\lambda_\omega-r) \tag{6.46}$$

combined with Eq. (6.32b). Consequently the real part of the electric susceptibility due to bound pairs can be estimated by (Ambegaokar *et al.*, 1978, 1980)

$$\mathrm{Re}\{\chi_B(\omega)\}=\chi_B(\lambda_\omega)\ \text{ for }\lambda_\omega<\lambda<\lambda_c=\infty\ . \tag{6.47}$$

Likewise the imaginary part of g can, in the expression for χ_B, be approximated by [compare Eq. (6.32b)]

$$\mathrm{Im}\{\tilde{g}(r,\omega)\}=\frac{\lambda_\omega^2r^2}{\lambda_\omega^4+r^4}\simeq\frac{\pi r}{4}\delta(r-\lambda_\omega)\ , \tag{6.48}$$

provided $\partial\chi_B(r)/\partial\ln(r)$ is slowly varying. Accordingly the imaginary part of $\chi_B(\omega)$ can be estimated by (Am-

begaokar *et al.*, 1978, 1980)

$$\mathrm{Im}\{\chi_B(\omega)\}=\frac{\pi}{4}\frac{\partial\chi_B(r)}{\partial\ln(r)}\bigg|_{r=\lambda_\omega} \quad \text{for } \lambda_\omega<\lambda<\lambda_c=\infty \ . \tag{6.49}$$

The complex dielectric constant $\varepsilon(\omega)$ [see Eq. (6.18)] can now be obtained by adding the contributions from the free and bound pairs given by Eqs. (6.26), (6.47), and (6.49). This is the main result of the reasoning by Ambegaokar *et al.*, (1978, 1980) and Ambegaokar and Teitel (1979).

The qualitative content of the resulting $\varepsilon(\omega)$ is the same as the alternative one described in Sec. VI.A.1. An application to the torsion pendulum experiment is described in Sec. IV.B.1.

3. Relation to vortex dynamics

In the description given above, the dynamics was introduced *ad hoc* into the Coulomb gas by postulating the Langevin equation (6.1). The object of this dynamical Coulomb gas model is to describe vortex fluctuations for superfluid-superconducting films. The Langevin equation (6.1) is consequently dictated by the requirement of being the Coulomb gas analogy of vortex dynamics (Ambegaokar *et al.*, 1978, 1980).

In this section we briefly motivate the Langevin equation (6.1) from vortex dynamics. A more complete discussion of vortex dynamics in the context of superfluid films can be found in Ambegaokar *et al.* (1980) and Trugman and Doniach (1982).

Consider the case of a single vortex on a superfluid film with position $\mathbf{r}(t)$ and vorticity s. An imposed superfluid mass flow $\mathbf{g}_{ex}$ gives rise to a force $\mathbf{F}$ acting on it, as given by (see Vinen, 1961, and Nozières and Vinen, 1966)

$$\mathbf{F}=s\frac{2\pi\hbar}{m^*}\left[\mathbf{g}_{ex}-\rho\frac{d\mathbf{r}(t)}{dt}\right]\times\hat{\mathbf{x}}_3 \ , \tag{6.50}$$

where $\hat{\mathbf{x}}_3$ is the unit vector perpendicular to the superfluid film. For a superconductor, this is the Lorentz force, which in the rest frame of the vortex reduces to [compare Eqs. (4.25) and (4.61)]

$$\mathbf{F}=s\frac{2\pi\hbar}{m^*}(\mathbf{g}_{ex}\times\hat{\mathbf{x}}_3)=s\varphi_0(\mathbf{j}_{ex}\times\hat{\mathbf{x}}_3) \ . \tag{6.51}$$

For a neutral superfluid, $\mathbf{F}$, given by Eq. (6.50) is usually referred to as the Magnus force.

A vortex moving relative to the substrate will experience additional forces. These are called the drag forces $\mathbf{F}_D$ and may be assumed to be of the form (see Vinen, 1961)

$$\mathbf{F}_D=\rho\left[-b\frac{d\mathbf{r}}{dt}-b's\hat{\mathbf{x}}_3\times\frac{d\mathbf{r}}{dt}\right] , \tag{6.52}$$

where b and b' are two constants. In case of a superconductor, there is often in addition the vortex pinning force $\mathbf{F}_p$ (see, for example, Huebner, 1979). The pinning force will, however, be neglected in the present description. As mentioned in Sec. V.B.2, flux pinning in many practical cases appears to limit the applicability of the Coulomb gas description of vortex fluctuations for temperatures lower than T_c.

A vortex has no inertia and consequently its motion is determined by the condition that the total force acting on it is equal to zero. Within the present description this means that

$$\mathbf{F}+\mathbf{F}_D=0 \tag{6.53}$$

with $\mathbf{F}$ and $\mathbf{F}_D$ given by Eqs. (6.50) and (6.52), respectively. The corresponding equation of motion is given by

$$\frac{dr}{dt}=\frac{s\,2\pi\hbar\mathcal{D}}{m^*T}\mathbf{g}_{ex}\times\hat{\mathbf{x}}_3+\mathcal{C}\mathbf{g}_{ex}+\boldsymbol{\eta} \ , \tag{6.54a}$$

where

$$\mathcal{D}=\frac{T}{\rho}\frac{b}{(2\pi\hbar/m^*-b')^2+b^2} \ , \tag{6.54b}$$

$$\mathcal{C}=\frac{1}{\rho}\frac{(2\pi\hbar/m^*-b')2\pi\hbar/m^*}{(2\pi\hbar/m^*-b')^2+b^2} \ , \tag{6.54c}$$

and $\boldsymbol{\eta}$ is the Langevin Gaussian noise term obeying Eq. (6.2) (Ambegaokar *et al.*, 1980).

For superconductors, the typical case for which the Coulomb gas analogy of vortex fluctuations should apply is that of thin "dirty" type-II superconducting films (Beasley *et al.*, 1979; see Sec. IV.B). In this case the $\mathcal{C}\mathbf{g}_{ex}$ term is very small compared to the first term on the right-hand side of Eq. (6.45a) and can be safely discarded (see, for example, Huebner, 1979). Consequently the equation of motion for a single vortex in this case reduces to

$$\frac{d\mathbf{r}}{dt}=\frac{s\mathcal{D}}{T^{CG}}\mathbf{D}+\boldsymbol{\eta} \ , \tag{6.55}$$

where $T^{CG}=T/[2\pi\rho(\hbar/m^*)^2]$ is the Coulomb gas temperature [see Eq. (4.20)] and $\mathbf{D}=m^*\mathbf{g}_{ex}\times\hat{\mathbf{x}}_3/(\hbar\rho)$ is the external electric field in the Coulomb gas analogy [compare Eq. (4.62) and Table IV]. The force acting on the vortex is modified by the presence of other vortices and is given by

$$\mathbf{F}^{\mathrm{eff}}=\frac{sm^*}{\hbar\rho}\mathbf{g}^{\mathrm{eff}}(\mathbf{r})\times\hat{\mathbf{x}}_3 \ , \tag{6.56}$$

where $\mathbf{g}^{\mathrm{eff}}(\mathbf{r})$ is the superfluid mass current felt by the vortex. The essence of the Coulomb gas analogy is that the force acting on a vortex can also be expressed as the "bare" force (i.e., the force in the absence of other vortices) plus the sum of the forces exerted on it by the other vortices (see Secs. IV.A and IV.B). Consequently, for the case considered, Eq. (6.1) is the Coulomb gas analogy of the Langevin equation for vortices where $\mathbf{F}_j^{\mathrm{eff}}(t)$ is the sum of the "bare" force $s_j\mathbf{D}(t)$ and the forces exerted on the Coulomb gas charge j by all the other Coulomb gas charges. Note that in this case no additional restriction is imposed on the Coulomb gas analogy of vortex fluctuations by the vortex dynamics. For example, the dynami-

cal Coulomb gas description given in Sec. IV.A.1 should also apply to the vortex fluctuations of a thin superconducting type-II film in a perpendicular magnetic field, i.e., the non-neutral Coulomb gas case. Such a case is described in Sec. VI.B.2.

In the case of a neutral superfluid, the $\mathcal{C}\mathbf{g}_{ex}$ term in Eq. (6.54) is *a priori* present. However, as is clear from Sec. VI.A.2, the dielectric constant $\varepsilon(\omega)$ only involves information on the sum $\sum_j s_j\langle \mathbf{r}_j(t)\rangle$, which for a *neutral* Coulomb gas (that is, one with an equal number of positive and negative charges) can be divided into pairs of particles, $\sum_j s_j\langle \mathbf{r}_{ij}\rangle = \sum_{\langle ij\rangle}\langle \mathbf{r}_{ij}\rangle$ [compare Eq. (6.21)]. The equation of motion for $\mathbf{r}_{ij}(t)$ involves only the effective force acting between the charges in the pair [compare Eqs. (6.36)]. The vortex analogy of this force is the force acting between two vortices with opposite vorticities. Obviously the $\mathcal{C}\mathbf{g}_{ex}$ term does not contribute to this force. As discussed above, the presence of other vortices changes $\mathbf{g}_{ex}$ to $\mathbf{g}^{\rm eff}(\mathbf{r})$. Thus, to the extent that the spatial variation of $\mathbf{g}^{\rm eff}$ can be neglected, the contribution arising from the $\mathcal{C}\mathbf{g}_{ex}$ term in Eq. (6.54a) can be discarded as far as the dielectric constant $\varepsilon(\omega)$ is concerned. In this sense the dynamical Coulomb gas model described in Sec. VI.A also applies to the neutral superfluid in a situation where there is no net vorticity (i.e., where there are equal numbers of vortices with vorticities $s=1$ and -1); Ambegaokar *et al.*, 1978, 1980).

The equation of motion given by Eq. (6.54) is valid in the rest frame of the substrate (Ambegaokar *et al.*, 1978, 1980). The torsion pendulum experiment for ^{4}He involves a moving substrate (see Secs. V.A and VI.B.1). A transformation to the laboratory frame is hence needed. However, the dielectric constant $\varepsilon(\omega)$ is a Galilean invariant quantity (Teitel, 1981). Consequently the dynamical Coulomb gas described in Sec. VI.A also covers this situation (Ambegaokar *et al.*, 1978, 1980).

B. Experimental examples

The aim of this section is to illustrate the connection between the "dynamical" version of the two-dimensional Coulomb gas described in Sec. VI.A and experiments reflecting the "dynamical" aspects of vortex fluctuations for ^{4}He films and superconducting films. Just as in Sec. V, dealing with experiments reflecting the "static" properties of vortex fluctuations, only a few examples out of the large body of work existing are selected and discussed. Again a subjective choice has been made with the criterion of illustrating the Coulomb gas aspects as clearly as possible.

The reader interested in a more complete description of the experimental work reflecting "dynamical" aspects is referred to review articles by Hebard and Fiory (1982) in the case of superconducting films and Glaberson and Donelly (1986) in the case of helium films.

Section VI.B.1 contains a discussion of the "dynamical" aspects of the torsion pendulum experiment for ^{4}He films. Some "static" aspects of this experiment were considered in Sec. V.A. The complex dielectric constant $\varepsilon(\omega)$ and its scaling properties measured for superconducting films in a perpendicular magnetic field are the subject of Sec. VI.B.2.

1. The torsion pendulum experiment for ^{4}He

The basic principle of the pendulum experiment for ^{4}He films, devised by Bishop and Reppy (1978), was sketched in Sec. V.A.

The dynamics involved in the experiment is related to the Coulomb gas dielectric constant $\varepsilon(\omega)$ in the following way: Let $X(t)$ denote the amplitude of the pendulum. The pendulum is driven at its resonance frequency ω, so that $X(t)=X_0{\rm Re}\{e^{-i\omega t}\}$. This pendulum dissipates energy, and this dissipation can be inferred by considering the case when the driving force suddenly ceases. The pendulum then becomes damped and $X(t)$ changes to $X(t)=X_0e^{-\tilde{\gamma}t}{\rm Re}\{e^{-i\omega t}\}$, where $\tilde{\gamma}$ is the relaxation constant. The equation of motion becomes (compare Sec. V.A)

$$(i\omega+\tilde{\gamma})^2\left[M-\frac{\Omega\rho}{\varepsilon(\omega)}\right]X_0+KX_0=0\ , \tag{6.57}$$

where K is the spring constant and M is the total oscillating mass including the helium. For $\tilde{\gamma}/\omega\ll 1$ and $\Omega\rho/M\ll 1$, the solution is

$$\omega=\left[{\rm Re}\left\{\frac{K}{M-\Omega\rho/\varepsilon(\omega)}\right\}\right]^{1/2}+\tilde{\gamma}^2$$
$$\simeq\left[\frac{K}{M}\right]^{1/2}\left[1+\frac{\Omega\rho}{2M}{\rm Re}\{1/\varepsilon(\omega)\}\right] \tag{6.58a}$$

and

$$\tilde{\gamma}=-\frac{1}{2\omega}{\rm Im}\left\{\frac{K}{M-\Omega\rho/\varepsilon(\omega)}\right\}$$
$$\simeq-\left[\frac{K}{M}\right]^{1/2}\frac{\Omega\rho}{2M}{\rm Im}\{1/\varepsilon(\omega)\}\ . \tag{6.58b}$$

From Eq. (6.58a) one obtains the period shift ΔP caused by the superfluidity,

$$\frac{\Delta P}{P}=\frac{\Omega\rho}{2M}{\rm Re}\{1/\varepsilon(\omega)\}=\frac{\Omega\rho_S}{2M}\ , \tag{6.59}$$

where ρ_S is the time-averaged superfluid density [compare Eqs. (5.1), (6.4), and (6.9)] and $P=2\pi\sqrt{M/K}$ is the period in the absence of superfluidity. The relaxation constant $\tilde{\gamma}$ is simply related to the friction force F_f experienced by the pendulum through $F_f\simeq-2\tilde{\gamma}M\omega X_0\sin(\omega t)$ [compare Eq. (6.57)]. The friction work during a period is consequently given by $2\pi\tilde{\gamma}\omega MX_0^2$, whereas the energy stored in the pendulum is $KX_0^2/2$. The dissipation in the experiment is measured by the quality factor Q, which may be defined as $2\pi Q^{-1}=$(dissipated energy per cycle)/(stored energy). For the pendulum this means that [compare Eq. (6.58b)]

$$\Delta Q^{-1}=2\tilde{\gamma}\sqrt{M/K}=-\frac{\Omega\rho}{M}\mathrm{Im}\{1/\varepsilon(\omega)\}\ , \qquad (6.60)$$

where ΔQ^{-1} is the change in Q^{-1} caused by the superfluidity. This change can be extracted from the experiment. Consequently, since the "bare" superfluid density ρ appears, to a good approximation, to be temperature independent for ^{4}He films, the torsion pendulum experiment should in a particularly direct way measure the temperature dependence of the real and imaginary parts for the Coulomb gas response function $1/\varepsilon(\omega)$.

A thorough account of the torsion pendulum experiment for ^{4}He films for the case when the ^{4}He is adsorbed on a Mylar substrate has been given by Agnolet (1983). Figure 31 shows a typical set of data obtained from such an experiment (Agnolet, 1983).

A comparison with Fig. 29 immediately suggests that a qualitative explanation of the measurements is given by vortex unbinding, as described by the dynamical version of the Coulomb gas model presented in Sec. VI.A.1. The general shape of the $\Delta P/P$ data and the ΔQ^{-1} data corresponds very well with the Coulomb gas quantities $\mathrm{Re}\{1/\varepsilon(\omega)\}$ and $-\mathrm{Im}\{1/\varepsilon(\omega)\}$, respectively. One notes that the dissipation peak is very close to the temperature where $\Delta P/P$ has half its T_c value, as predicted by the approximate theory described in Sec. VI.A.1 ($T_c\simeq 514$ mK is determined as described in Sec. V.A.1; compare Fig. 19).

A more quantitative test of the approximate description provided by Eqs. (6.15) is represented by the dashed curve in Fig. 31. This test is constructed in the following way:

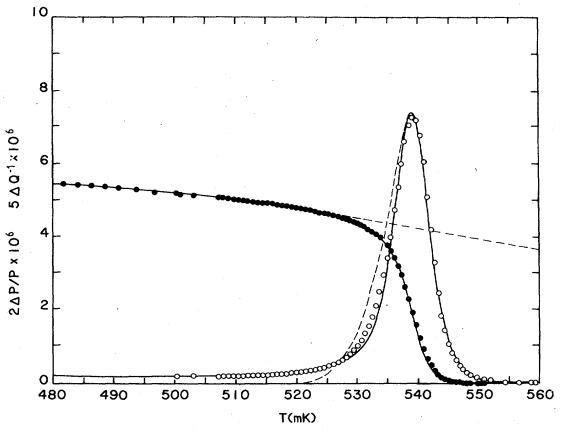

FIG. 31. Example of results obtained from the torsion pendulum experiment for ^{4}He adsorbed on a Mylar substrate (from Agnolet, 1983): ●, the measured period shift ΔP; ○, the measured change in the inverse quality factor ΔQ^{-1}. The dashed straight line represents a rough estimate of the dielectric constant $\tilde{\varepsilon}$ (see text). The dashed curve is obtained from the Coulomb gas prediction given by Eqs. (6.15) with *one* adjustable parameter (see text). It coincides with the solid curve on the high-temperature side of the peak. The two solid curves are fits to the Coulomb gas description by Ambegaokar *et al.* (1978, 1980) and Ambegaokar and Teitel (1979), which together involve *six* adjustable parameters (from Agnolet, 1983; see text).

$\mathrm{Re}\{1/\varepsilon(\omega)\}$ is determined as the ratio $\Delta P(T)/\Delta P(0)$ (compare Sec. V.A.2). The length λ_c is assumed to be infinite and the λ_ω/λ is determined from the functional form of $\mathrm{Re}\{1/\varepsilon(\omega)\}$ given by Eq. (6.15a). The dielectric function $\tilde{\varepsilon}$ is determined from the dashed straight line in Fig. 31, which just represents a linear extrapolation from temperatures somewhat below T_c, providing a slight improvement of the rough estimate $\tilde{\varepsilon}=1$ used in Fig. 29. $\mathrm{Im}\{1/\varepsilon(\omega)\}$ is then obtained for each temperature from Eq. (6.15b) using $\tilde{\varepsilon}$ and λ_ω/λ as input.

Finally the scale factor $\Delta Q^{-1}/[-\mathrm{Im}\{1/\varepsilon(\omega)\}]$ in Eq. (6.60) is determined by adjusting the calculated $-\mathrm{Im}\{1/\varepsilon(\omega)\}$ to the measured peak height of ΔQ^{-1}. The agreement between the Coulomb gas description given by Eqs. (6.15) and the measured ΔQ^{-1} is striking, especially considering that only *one* adjustable parameter is involved, i.e., the scale factor. In the temperature interval from the ΔQ^{-1} peak and upwards, the agreement is excellent (the dashed curve coincides with the solid curve in Fig. 31 for temperatures above the peak). This interval is also very insensitive to the rough estimate of $\tilde{\varepsilon}$. The small deviation on the low-temperature side down to T_c can be explained within the Coulomb gas description as being caused by a too crude estimate of $\tilde{\varepsilon}$. The observed finite ΔQ^{-1} below T_c can be explained within the Coulomb gas model by a finite λ_c (see Fig. 30). Consequently the Coulomb gas model gives a good qualitative account of the observed features for $\Delta P/P$ and ΔQ^{-1}. In addition Eqs. (6.15) give a surprisingly good quantitative account of the measurements.

The solid curves in Fig. 31 are fits to the theory by Ambegaokar *et al.* (1978, 1980) and Ambegaokar and Teitel (1979), which was described in Sec. VI.A.2 (from Agnolet, 1983). The details of the fitting procedure can be found in Bishop and Reppy (1980) and Teitel (1981). The functional dependencies are obtained from Kosterlitz equations [Eqs. (3.16)]. The procedure involves *six* adjustable parameters, *five* of which are contained in the theory and *one* added *ad hoc* in order to account for the finite ΔQ^{-1} below T_c. As seen in Fig. 31, this gives an excellent fit to the measurements. However, some caution may be advisable when assessing the significance of this agreement. The approximate theory described in Sec. VI.A.2 is *a priori* only expected to give a *qualitative* account of the "dynamical" Coulomb gas model. Six free parameters in a qualitatively correct theory can certainly provide a very good fit, even if the details of the theory are quite oversimplified. The significance of the agreement has then to be judged according to the physical plausibility of the obtained fitting parameters. As pointed out by Agnolet (1983), the plausibility of the fitting parameters differs for various ^{4}He films. This suggests that the theory is, in fact, somewhat oversimplified, in spite of the good agreement. At any rate, the version of the "dynamical" Coulomb gas theory by Ambegaokar *et al.*, (1978, 1980) and Ambegaokar and Teitel (1979) (outlined in Sec. V.A.2) gives a convincing qualitative account of the experimental results.

In summary, the two versions of the "dynamical"

Coulomb gas model given in Secs. VI.A.1 and VI.A.2 both strongly support the interpretation of the $\Delta P/P$ data and the ΔQ^{-1} data for the torsion pendulum experiment for ^{4}He films on Mylar substrates in terms of vortex unbinding.

2. The Coulomb gas dielectric function for superconducting films

The Coulomb gas dielectric function $\varepsilon(\omega)$ for superconducting films can be rather directly extracted from an experiment devised by Hebard and Fiory (1980). In this experiment the superconduction film is placed between a driver coil and a receiver coil. The driver coil produces an electric field $\mathbf{E}(t)$ acting on the film. A current $\mathbf{j}_{\rm tot}(t)$ is induced on the film and the induction produced by this current on the receiver coil is measured. From this measurement one obtains the complex impedance $Z(\omega)$ defined by

$$\mathbf{E}(\omega)=Z(\omega)\mathbf{j}_{\rm tot}(\omega)\ . \tag{6.61}$$

In the absence of vortices, the relation between the electric field and the current can be obtained from the Josephson relation (Josephson, 1962) $\Delta V=\hbar\Delta\theta/e^*$ relating the voltage drop ΔV between two points to the corresponding difference in the order-parameter phase $\Delta\theta$. This leads to the relation [Halperin and Nelson, 1979; compare Eqs. (4.23) and (4.25)]

$$\mathbf{E}(\omega)=-i\omega L_k\mathbf{j}(\omega)\ , \tag{6.62a}$$

where

$$L_k=\left[\frac{m^*}{e^*}\right]^2\Big/\rho \tag{6.62b}$$

is the sheet kinetic inductance. The presence of vortices modifies the current to $\mathbf{j}_{\rm tot}(\omega)=\mathbf{j}(\omega)/\varepsilon(\omega)$, which follows from Eq. (6.6) since $\mathbf{E}^{\rm eff}(\omega)=\mathbf{D}(\omega)/\varepsilon(\omega)$ by definition. Consequently the measured complex impedance is directly related to the Coulomb gas dielectric constant by

$$Z(\omega)=-i\omega L_k\varepsilon(\omega)\ . \tag{6.63}$$

Figure 32 shows the result of such an experimental determination of the Coulomb gas dielectric constant $\varepsilon(\omega)$ (from Fiory and Hebard, 1982). A thin granular aluminum film was used and the data were taken as a function of a perpendicular magnetic field $B_{\rm ex}$ at a constant temperature somewhat below the experimentally determined Kosterlitz-Thouless temperature. The eight solid curves in Fig. 32 represent the real and imaginary parts of $\varepsilon(\omega)$ for four different frequencies spanning an interval of three orders of magnitude from 2 to 2000 (kHz). The dielectric constant $\varepsilon(\omega)$ is plotted in the figure as a function of $B_{\rm ex}/\omega$.

From the point of view of Coulomb gas theory, the experimentally determined $\varepsilon(\omega)$ shown in Fig. 32 corresponds to a non-neutral Coulomb gas at a fixed temperature $T^{\rm CG}$ below $T_c^{\rm CG}$. The screening length is consequently related to $B_{\rm ex}$ in a simple way [compare Eqs. (2.21b), (2.44), and (4.38)],

$$\lambda^{-2}=\frac{2\pi n_F}{\tilde{\varepsilon}T^{\rm CG}}=\frac{2\pi\langle\Delta n\rangle}{\tilde{\varepsilon}T^{\rm CG}}=\frac{2\pi B_{\rm ex}}{\tilde{\varepsilon}T^{\rm CG}\varphi_0}\ . \tag{6.64}$$

The "dynamical" version of the Coulomb gas model given in Sec. VI.A.1 provides an approximate description. Consequently $\varepsilon(\omega)$ is approximately given by Eqs. (6.15) with $(\lambda_\omega/\lambda)^2\sim B_{\rm ex}/\omega$, since $\lambda_\omega\sim\sqrt{\mathcal{D}/\omega}$ [see below Eq. (6.12)]. The explicit expressions for $\mathrm{Re}\{\varepsilon(\omega)\}$ and $\mathrm{Im}\{\varepsilon(\omega)\}$, as obtained from Eqs. (6.15) for this special case, are

$$\mathrm{Re}\{\varepsilon(\omega)\}=\tilde{\varepsilon}\frac{(1-\tilde{\mathcal{C}}B_{\rm ex}/\omega)[1-(\tilde{\mathcal{C}}B_{\rm ex}/\omega)^2]}{(1-\tilde{\mathcal{C}}B_{\rm ex}/\omega)^2+[2\tilde{\mathcal{C}}B_{\rm ex}\ln(\tilde{\mathcal{C}}B_{\rm ex}/\omega)/\pi\omega]^2} \tag{6.65a}$$

and

$$\mathrm{Im}\{\varepsilon(\omega)\}=\tilde{\varepsilon}\frac{[2\tilde{\mathcal{C}}B_{\rm ex}\ln(\tilde{\mathcal{C}}B_{\rm ex}/\omega)/\pi\omega][1-(\tilde{\mathcal{C}}B_{\rm ex}/\omega)^2]}{(1-\tilde{\mathcal{C}}B_{\rm ex}/\omega)^2+[2\tilde{\mathcal{C}}B_{\rm ex}\ln(\tilde{\mathcal{C}}B_{\rm ex}/\omega)/\pi\omega]^2}\ , \tag{6.65b}$$

where $\tilde{\mathcal{C}}$ is a constant proportional to the diffusion constant $\mathcal{D}$.

A striking consequence of the Coulomb gas description is the prediction that the dielectric constant should be a function only of the variable $B_{\rm ex}/\omega$ [compare Eqs. (6.65)]. This scaling relation for $\varepsilon(\omega)$ is of the same type as the scaling relations discussed in Sec. IV.D.2. It follows since the dimensionless quantity $\varepsilon(\omega)$ can at most be a function of the two dimensionless variables $T^{\rm CG}$ and λ/λ_ω. Consequently the existence of this scaling relation may be expected to be rather independent of the explicit model assumptions.

As can be seen from Fig. 32, the predicted $B_{\rm ex}/\omega$ scaling is borne out to an impressively high degree for $\mathrm{Im}\{\varepsilon(\omega)\}$ (Fiory and Hebard, 1982). It is borne out to a somewhat lesser degree for $\mathrm{Re}\{\varepsilon(\omega)\}$, and this deviation from scaling could be attributed to vortex pinning (Fiory and Hebard, 1982).

A more quantitative test of the approximate description provided by Eqs. (6.65) is represented by the dashed curves in Fig. 32. As can be seen in the figure, the theoretical $\mathrm{Im}\{\varepsilon(\omega)\}$ curve agrees very well with the measured four $\mathrm{Im}\{\varepsilon(\omega)\}$ curves. In fact this agreement is quite impressive, considering that Eq. (6.65b) contains only *one* free parameter, i.e., the constant $\tilde{\mathcal{C}}$, and that the agreement involves five orders of magnitude in $B_{\rm ex}/\omega$, three orders of magnitude in $\mathrm{Im}\{\varepsilon(\omega)\}$, and three orders of magnitude in ω.

The constant $\tilde{\mathcal{C}}$ can be approximately related to the Ginzburg-Landau coherence length ξ in the following way: Eq. (6.64) and the relation $(\lambda_\omega/\lambda)^2=\tilde{\mathcal{C}}B_{\rm ex}/\omega$ (which is the definition of $\tilde{\mathcal{C}}$), together with the approximations $\lambda_\omega\simeq\sqrt{\mathcal{D}/\omega}$ and $\tilde{\varepsilon}T^{\rm CG}\simeq\frac{1}{4}$, give $\tilde{\mathcal{C}}\simeq 8\pi\mathcal{D}/\varphi_0$. The diffusion constant is related to the vortex mobility by $\tilde{\mu}=\mathcal{D}/k_BT$, and $\tilde{\mu}$ is related to the Ginzburg-Landau coherence length ξ by the Bardeen-Stephen relation (Bar-

deen and Stephen, 1965), $\tilde{\mu}=(\xi e^*)^2/(2\hbar^2\pi\sigma_N)$ where σ_N is the normal-state conductivity. By this sequence of relations one finds that the theoretical $\mathrm{Im}\{\varepsilon(\omega)\}$ curve given in Fig. 32 corresponds to $\xi\simeq 20$ (Å), which is indeed a physically rather plausible magnitude of the Ginzburg-Landau coherence length for a granular aluminum film well below T_c.

The theoretical $\mathrm{Re}\{\varepsilon(\omega)\}$ curve is obtained from Eq. (6.65a) with the same value of $\tilde{C}$ as the $\mathrm{Im}\{\varepsilon(\omega)\}$ curves. Consequently the theoretical $\mathrm{Re}\{\varepsilon(\omega)\}$ curve involves *no* adjustable parameter. Comparison with the four measured Re $\{\varepsilon(\omega)\}$ curves in Fig. 32 shows that the theoretical curve predicted from Coulomb gas theory appears to have the correct magnitude and functional form.

In summary, the dielectric constant $\varepsilon(\omega)$ measured for superconducting films appears to agree very well with the Coulomb gas model predictions. In particular, the B_{ex}/ω scaling for $\mathrm{Im}\{\varepsilon(\omega)\}$ is borne out to a high degree (Fiory and Hebard, 1982) and the approximate functional form of $\mathrm{Im}\{\varepsilon(\omega)\}$ predicted from the Coulomb gas model agrees very well with the measurements.

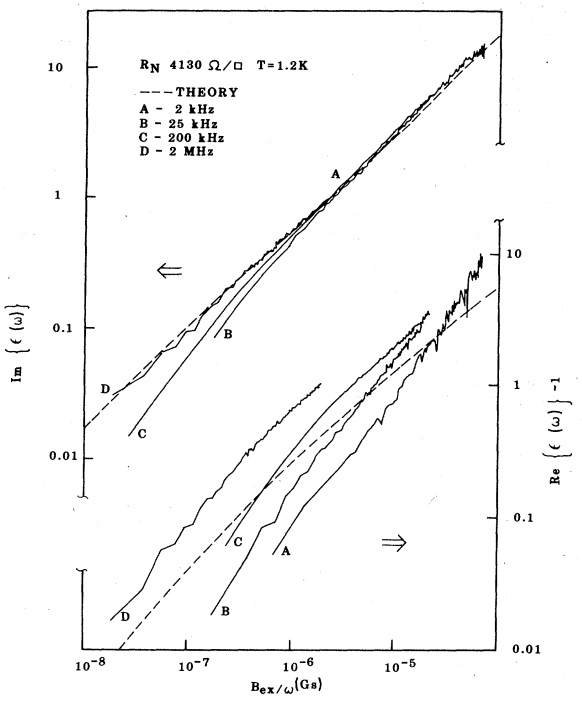

FIG. 32. The dielectric constant $\varepsilon(\omega)$ for a superconducting film (from Fiory and Hebard, 1982). The measurements shown are for a granular aluminum film with normal-state resistance $R_N=4130$ (Ohms per square) taken as a function of an external perpendicular magnetic field B_{ex} at a fixed temperature $T=1.2$ (K) well below the experimentally determined Kosterlitz-Thouless temperature. The eight solid curves represent the real and imaginary parts of $\varepsilon(\omega)$ plotted as functions of B_{ex}/ω for four different frequencies, 2, 25, 200, and 2000 kHz, respectively, as indicated in the figure. The two dashed curves are the Coulomb gas predictions given by Eqs. (6.65) involving *one* adjustable parameter (see text).

C. Vortex dynamics and Maxwell's equations

There exists an interesting analogy between Maxwell's equations for the two-dimensional Coulomb gas and vortex renormalization of third-sound propagation for helium films (Ambegaokar *et al.*, 1980). This analogy will be reviewed in two steps. First, the analogy between a two-dimensional membrane with vortices in the displacement field and Maxwell's equations is described (Sec. VI.C.1). Then it is explained how this model of a two-dimensional membrane can be viewed as a simple model for ^{4}He films, and the qualitative Coulomb gas predictions for the third sound are discussed (Sec. VI.C.2).

1. Maxwell's equations and a two-dimensional membrane

In this section we consider the model specified by the Hamiltonian

$$H_M=\int d^2r\left[\frac{\tilde{\sigma}}{2}\left[\frac{\partial\varphi}{\partial t}(\mathbf{r},t)\right]^2+\frac{\tau}{2}[\nabla\phi(\mathbf{r},t)]^2\right], \tag{6.66}$$

which can be interpreted as the Hamiltonian for a two-dimensional membrane with displacement field $\varphi(\mathbf{r},t)$, mass density $\tilde{\sigma}$, and tension τ. The Euler-Lagrange equation for this model is the usual wave equation

$$\frac{1}{c_0^2}\frac{\partial^2\varphi}{\partial t^2}=\nabla^2\varphi\,, \tag{6.67}$$

where $c_0=\sqrt{\tau/\tilde{\sigma}}$ is a characteristic velocity. If ϕ is a "smooth" field, the dispersion relation for waves is

$$\omega=c_0k\,, \tag{6.68}$$

where c_0 is the wave velocity and k the absolute value of the wave vector.

The object is to describe how the dispersion relation (6.68) is modified when the field φ contains vortex singularities of the same type as discussed in Sec. IV.A. The field φ may in this case be expressed as the sum of two parts, $\varphi=\varphi_R+\varphi_S$, where φ_R is a "smooth" field and φ_S contains the vortex singularities. In analogy with Eq. (4.5), a vortex singularity is defined as a point around which the line integral $\int\mathbf{ds}\cdot\nabla\varphi$ is $\pm 2\pi$. These vortex singularities can, as discussed in Sec. IV.A.1, be specified in terms of a scalar potential W [compare Eqs. (4.4b) and (4.6)].

$$\nabla\varphi_S=\nabla\times\hat{\mathbf{x}}_3W(\mathbf{r},t)\,, \tag{6.69a}$$

where

$$\nabla^2W(\mathbf{r},t)=-\sum_i 2\pi s_i\delta(\mathbf{r}-\mathbf{r}_i) \tag{6.69b}$$

and $\mathbf{r}_i(t)$ ($s_i=\pm 1$) denotes the position (vorticity) of vortex i. The vortex positions are assumed to be functions of time. From Eqs. (6.69) the explicit form of the field φ_S may be obtained,

$$\varphi_S=\sum_i s_i \arctan\left[\frac{x_2-x_2^i}{x_1-x_1^i}\right], \tag{6.70}$$

where $\mathbf{r}_i=(x_1^i,x_2^i)$.

Precisely as in Sec. IV.A.2 the vortices can be interpreted as Coulomb gas charges. The Coulomb gas charge density $\Delta n(\mathbf{r},t)$ is in the present case given by

$$\Delta n(\mathbf{r},t)=\sum_i s_i \delta(\mathbf{r}-\mathbf{r}_i), \tag{6.71}$$

and the current density $j_C(\mathbf{r},t)$ is associated with the Coulomb gas charges by

$$\mathbf{j}_C(\mathbf{r},t)=\sum_i s_i\left[\frac{d}{dt}\mathbf{r}_i\right]\delta(\mathbf{r}-\mathbf{r}_i). \tag{6.72}$$

The total field φ obeys a set of differential equations that can be straightforwardly inferred (Côté and Griffin, 1986). The first is the Euler-Lagrange equation (6.67), which may be cast into the form

$$\nabla\times(\nabla\varphi\times\hat{\mathbf{x}}_3)=-\hat{\mathbf{x}}_3\nabla^2\varphi=-\frac{1}{c_0^2}\frac{\partial^2}{\partial t^2}\varphi\hat{\mathbf{x}}_3. \tag{6.73a}$$

The second is Eq. (6.69), which may be rewritten as

$$\nabla\cdot(\nabla\varphi\times\hat{\mathbf{x}}_3)=\hat{\mathbf{x}}_3\cdot(\nabla\times\nabla\varphi)=2\pi\Delta n(\mathbf{r}). \tag{6.73b}$$

The third can be directly inferred from Eqs. (6.69), (6.70), and the definition (6.72), i.e., (Côté and Griffin, 1986),

$$\frac{\partial}{\partial t}\nabla\varphi_S=\nabla\left[\frac{\partial}{\partial t}\varphi_S\right]-2\pi\hat{\mathbf{x}}_3\times\mathbf{j}_C,$$

which may be recast into the form

$$\nabla\times\left[\frac{\partial}{\partial t}\varphi\hat{\mathbf{x}}_3\right]=-\hat{\mathbf{x}}_3\times\nabla\left[\frac{\partial}{\partial t}\varphi\right]$$
$$=\frac{\partial}{\partial t}(\nabla\varphi\times\hat{\mathbf{x}}_3)+2\pi\mathbf{j}_C. \tag{6.73c}$$

This set of equations can be expressed in terms of two coupled vector fields $\mathbf{E}$ and $\mathbf{B}$ defined by

$$\mathbf{E}=\nabla\varphi\times\hat{\mathbf{x}}_3, \tag{6.74a}$$

$$\mathbf{B}=\frac{1}{c_0}\left[\frac{\partial\varphi}{\partial t}\right]\hat{\mathbf{x}}_3. \tag{6.74b}$$

In terms of these fields, Eqs. (6.73a)–(6.73c) turn into, respectively,

$$\nabla\times\mathbf{E}=-\frac{1}{c_0}\frac{\partial}{\partial t}\mathbf{B}, \tag{6.75a}$$

$$\nabla\cdot\mathbf{E}=2\pi\Delta n, \tag{6.75b}$$

$$\nabla\times\mathbf{B}=\frac{1}{c_0}\frac{\partial}{\partial t}\mathbf{E}+\frac{2\pi}{c_0}\mathbf{j}_C, \tag{6.75c}$$

and

$$\nabla\cdot\mathbf{B}=0, \tag{6.75d}$$

where the last one follows directly from the definition (6.74b). Equations (6.75) may be recognized as Maxwell's equations with $\mathbf{E}$ and $\mathbf{B}$ identified as the electric and magnetic fields, respectively. Since all spatial dependencies are in two-dimensional space spanned by $\mathbf{r}=(x_1,x_2)$, Eqs. (6.75) are equivalent to Maxwell's equations for line charges along the $\hat{\mathbf{x}}_3$ axis. The formal analogy between vortices and Maxwell's equations in two dimensions was pointed out by Ambegaokar *et al.* (1980). The version described here is mainly due to Côté and Griffin (1986).

The solutions of Maxwell's equations are, of course, well known. For an isotropic system there are two possible wave modes. The one of interest in the present context is associated with the transverse part of $\mathbf{E}^T$ and is given by [compare Eqs. (6.75a) and (6.75c)]

$$\left[k^2-\frac{\omega^2}{c_0^2}\varepsilon^T(k,\omega)\right]\mathbf{E}^T(k,\omega)=0, \tag{6.76}$$

where $\varepsilon^T(k,\omega)$ is the transverse part of the dielectric response function defined by

$$\mathbf{E}^T(k,\omega)=\mathbf{E}^T_{\rm ex}(k,\omega)/\varepsilon^T(k,\omega) \tag{6.77}$$

and $\mathbf{E}^T_{\rm ex}$ is the transverse part of a small external electric field. In the small-k limit, $\varepsilon^T(k,\omega)$ reduces to $\varepsilon(\omega)$, where $\varepsilon(\omega)$ corresponds to the Coulomb gas dielectric constant introduced in Sec. VI.A. The dispersion relation (6.76) gives the modification of Eq. (6.68), caused by vortices, in terms of a Coulomb gas dielectric response function. In the limit of small k this modification can be expressed as follows: The wave velocity is renormalized [for $\mathrm{Im}\{\varepsilon(\omega)\}\ll\mathrm{Re}\{\varepsilon(\omega)\}$] to

$$c=c_0/\sqrt{\mathrm{Re}\{\varepsilon(\omega)\}} \tag{6.78a}$$

and the wave becomes damped. This damping can be expressed in terms of the quality factor Q [defined in connection with Eq. (6.60)], which is proportional to the ratio between the real and imaginary parts of the frequency, i.e.,

$$Q^{-1}=\frac{2\,\mathrm{Im}\{\omega\}}{\mathrm{Re}\{\omega\}}=\frac{\mathrm{Im}\{\varepsilon(\omega)\}}{\mathrm{Re}\{\varepsilon(\omega)\}}, \tag{6.78b}$$

where the last equality holds in the small $\mathrm{Im}\{\varepsilon(\omega)\}/\mathrm{Re}\{\varepsilon(\omega)\}$ limit.

The second possible wave mode is just the plasmon oscillation defined by $\varepsilon^L(k,\omega)=0$, where ε^L is the longitudinal part of the dielectric response function.

In order to form a complete description, we need also to specify a dynamical equation for the vortices. In the next section we describe how the present model supplemented with the Langevin equation (6.1) turns into a simple model of third-sound propagation for ^{4}He films.

2. ^{4}He films and third sound

The helium atoms are attracted to the substrate by van der Waal's interaction. A simple model of a helium film is obtained by assuming that the total energy of the film

is the sum of the kinetic energy due to the superfluid mass current [see Eq. (4.8)] and the potential energy Υ due to the substrate interaction, i.e.,

$$H_{\rm He}=\int d^2r\left[\frac{\mathbf{g}^2(\mathbf{r},t)}{2\rho}+\Upsilon(\mathbf{r},t)\right]. \tag{6.79}$$

Since the van der Waal's attraction depends only on the distance from the substrate surface, it can be assumed that Υ depends only on the thickness of the film, $d(\mathbf{r},t)$. Consequently, to lowest order in the deviation from the average thickness d_0, the potential Υ is given by

$$\Upsilon(\mathbf{r},t)=\frac{K}{2}[d(\mathbf{r},t)-d_0]^2\,, \tag{6.80}$$

where K is a constant (and Υ is measured relative to the equilibrium potential).

The potential Υ gives rise to a force density $-\nabla\Upsilon(\mathbf{r})$ acting on the helium, and this force is coupled to the superfluid mass current by

$$\frac{\partial}{\partial t}\mathbf{g}(\mathbf{r},t)=-\nabla\Upsilon(\mathbf{r},t)=-K\nabla[d(\mathbf{r},t)-d_0]\,. \tag{6.81}$$

This is just Newton's second law for the helium mass element at $\mathbf{r}$, since the motion of the normal part of the fluid relative to the substrate can be completely neglected in comparison with the superfluid part.

Let us represent $\mathbf{g}$ by a scalar field φ, so that

$$\mathbf{g}(\mathbf{r},t)=\nabla\varphi(\mathbf{r},t)\,. \tag{6.82}$$

It then follows from Eq. (6.81) that

$$\frac{\partial}{\partial t}\varphi(\mathbf{r},t)=-K[d(\mathbf{r},t)-d_0] \tag{6.83}$$

(up to an arbitrary function of time, which can be absorbed into the definition of φ). Consequently the model Hamiltonian $H_{\rm He}$ can be written [compare Eq. (6.79)]

$$H_{\rm He}=\int d^2r\left[\frac{[\nabla\varphi(\mathbf{r},t)]^2}{2\rho}+\frac{1}{2K}\left[\frac{\partial\varphi}{\partial t}(\mathbf{r},t)\right]^2\right], \tag{6.84}$$

which may be recognized as the Hamiltonian for the two-dimensional membrane discussed in Sec. VI.C.1 with the identifications $\tau=1/\rho$ and $\tilde{\sigma}=1/K$ [see Eq. (6.66)].

The inclusion of vortices in the two-dimensional membrane model described in Sec. VI.C.1 corresponds precisely to vortices in the superfluid mass current g [compare Eqs. (4.4b), (4.5), (6.69), and (6.82)], and the electric field $\mathbf{E}$ corresponds, up to a constant factor, to the electric field in the superfluid—Coulomb gas analogy [compare Eq. (6.55)]. In addition, the vortices in the superfluid obey the Langevin equation (6.1) (see Sec. VI.B.3).

Thus all the results from the preceding section apply to this simple description of helium films. The wave mode given by Eq. (6.76) corresponds in the helium case to a surface wave called third sound. The main result for third sound obtained from the present simple model is the renormalization of the sound velocity and the damping due to vortices given by Eqs. (6.78), where $\varepsilon(\omega)$ is the Coulomb gas dielectric constant discussed in Sec. VI.A. These results are due to Ambegaokar *et al.* (1980).

The Coulomb gas dielectric constant $\varepsilon(\omega)$ was discussed in Secs. VI.A.1 and VI.A.2. Its predominant feature is that the real and imaginary parts of $\varepsilon(\omega)$ increase rapidly as the Coulomb gas temperature $T^{\rm CG}$ is raised above the Kosterlitz-Thouless temperature $T_c^{\rm CG}$. This is a consequence of the rapidly decreasing screening length λ due to charge unbinding [compare Eqs. (6.15)—(6.17)]. The Coulomb gas temperature is proportional to T/ρ, where T is the real temperature and ρ is the "bare" areal superfluid density [see Eq. (4.20)]. For a fixed temperature the "bare" areal superfluid density ("real" density times film thickness) is expected to be at least roughly proportional to the film thickness d_0. Consequently $T^{\rm CG}$ increases with decreasing film thickness for fixed T. For the third-sound predictions given by Eqs. (6.78), this means that the Kosterlitz-Thouless transition should be reflected in a rapid decrease in the sound velocity and the quality factor Q with decreasing d_0. This qualitative prediction is borne out by experiment (Rudnick, 1978). Figure 33 gives an illustration for helium adsorbed on an Al_2O_3 powder with powder size of 1 μm (From Kotsubo and Williams, 1986). From the point

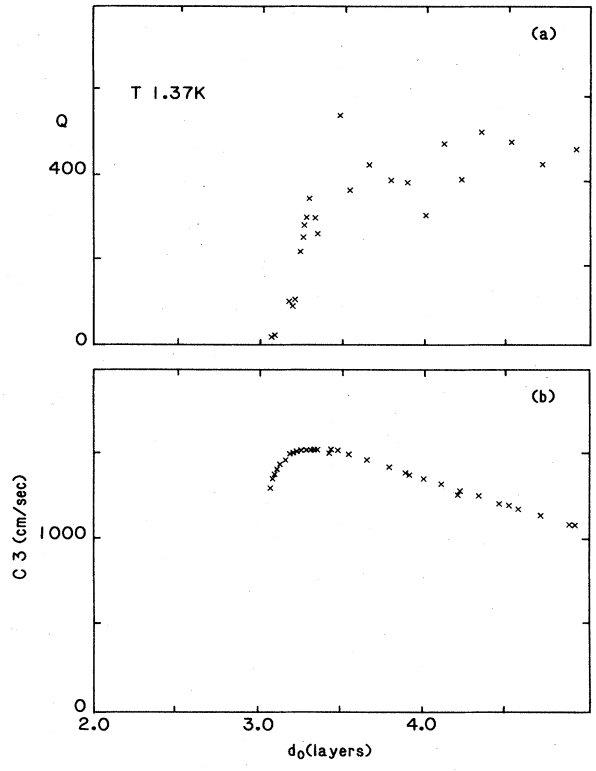

FIG. 33. Measured Q value (a) and velocity c_3 (b) for third-sound propagation as a function of helium film thickness d_0 in the case of ^{4}He adsorbed on an Al_2O_3 powder of powder size 1 μm (from Kotsubo and Williams, 1986). The Coulomb gas prediction is that the Q value and the velocity c_3 should rapidly decrease with decreasing thickness d_0 at the Kosterlitz-Thouless transition. These predictions are in accord with the data.

of view of the Coulomb gas model, this corresponds to the finite-λ_c-case, which should give a somewhat broader transition than the $\lambda_c = \infty$ case (Kotsubo and Williams, 1986; see Sec. VI.A.1).

A more realistic description of third-sound propagation for helium films involves thermal gradients (because the wave is not isothermal) and mass transport through the helium vapor above the film (Bergman, 1969, 1971). The inclusion of vortices in the description of third sound is due to Ambegaokar *et al.* (1980). A thorough discussion of vortices in connection with third sound has been given by Teitel (1982).

By starting from a more detailed description of third sound it is also possible to test the universal jump prediction by Nelson and Kosterlitz (1977) given by Eq. (4.44), and good agreement with this prediction has been found (Rudnick, 1978; see also Bishop and Reppy, 1980).

VII. MISCELLANEOUS COMMENTS

In the preceding sections the two-dimensional Coulomb gas *per se* and its connection to vortex fluctuations for superfluid-superconducting films have been described in some detail. The object has been to provide a useful perspective on the subject of vortex fluctuations. The purpose of this section is to broaden somewhat, by miscellaneous comments, the admittedly rather narrow perspective presented so far.

A. ^{4}He films

1. Variants of the torsion pendulum experiment

The torsion pendulum experiment with ^{4}He adsorbed on porous Vycor glass gives result somewhat different from the Mylar substrate case described in Sec. VI.B.1 (Berthold *et al.*, 1977; Smith *et al.*, 1978; Bishop *et al.*, 1981; Crooker *et al.*, 1983): the ΔP transition is broader and no ΔQ^{-1} dissipation is observed.

It has been pointed out by Kotsubo and Williams (1984, 1986) that the finite size of the Vycor grains—from the Coulomb gas point of view—implies a Coulomb gas with a *finite* λ_c, where λ_c is related to the grain size. As is obvious from Fig. 30, such a finite λ_c indeed accounts for the observed broader ΔP transition and a nonobservable ΔQ^{-1}. Kotsubo and Williams (1986) have made a systematic study of the transition from ^{4}He adsorbed in Al_2O_3 packed powders with decreasing powder size, using third-sound techniques, and the results to some extent support this qualitative Coulomb gas explanation.

Other possible explanations of the Vycor glass case involve vortex-pinning effects (Browne and Doniach, 1982; Yu, 1982) or indeed no vortices at all (Crooker *et al.*, 1982; Reppy, 1984). In this latter case the transition for low coverages should correspond to that of a dilute three-dimensional Bose fluid (Rasolt *et al.*, 1984; Weichman *et al.*, 1986).

One interesting extension of the torsion pendulum experiment for ^{4}He adsorbed on Mylar substrates, from the Coulomb gas point of view, is to perform the experiment in a rotating frame (Adams and Glaberson, 1986). As discussed in Sec. IV.A.2, a rotating frame with rotation velocity $\omega^{\rm rot}$ corresponds to a non-neutral Coulomb gas with $\langle \Delta n \rangle = \omega^{\rm rot} m^*/(\pi\hbar)$ [see Eq. (4.21)]. The properties of the "dynamical" Coulomb gas, as given by Eqs. (2.44) and (6.15b), do indeed provide a qualitative explanation of the observed $\omega^{\rm rot}$ dependence of the ΔQ^{-1} data. Adams and Glaberson (1986), on the other hand, have proposed an alternative explanation of their ΔQ^{-1} data in terms of a diverging diffusion constant $\mathcal{D}$.

2. Thermal conductance

The thermal conductance for a ^{4}He film should also, owing to the presence of vortices, reflect Coulomb gas properties. Ambegaokar *et al.* (1980) have shown that the thermal conductance, when expressed in the Couloumb gas analogy, should be inversely proportional to the density of free Coulomb gas charges n_F. Teitel (1982) gives a detailed discussion. Since the thermal conductance is inversely proportional to the density of free Coulomb gas charges, it should reflect the Kosterlitz-Thouless transition through a rapid decrease with increasing Coulomb gas temperature $T^{\rm CG} \sim T/\rho$. This prediction is in good agreement with measurements (Agnolet *et al.*, 1981; Maps and Hallock, 1981; Finotello and Gasparini, 1985). The n_F dependence on an external electric field D (compare Secs. II.B.3 and V.B.2) can also be extracted from measurements of the thermal conductance and is in accord with the Coulomb gas predictions (Joseph and Gasparini, 1982; Maps and Hallock, 1982).

B. Superconducting films

1. Thermoelectric vortex effect

A thermoelectric vortex effect for thin superconducting films has been observed by Lee *et al.* (1985) close to the experimentally determined Kosterlitz-Thouless transition. An explanation in terms of a vortex model involving a somewhat *ad hoc* vortex-relaxation process has been proposed by Garland and VanHarlingen (1985). The plausibility of this model has, however, been questioned by Gray (1986), who proposes a different edge vortex explanation.

From the point of view of the Coulomb gas analogy discussed in this article, it is interesting to note that the thermoelectric experiment by Lee *et al.* (1985) has a direct Coulomb gas interpretation and that the Coulomb gas model by itself does provide a possible qualitative explanation without any added *ad hoc* assumptions.

The experiment (Lee *et al.*, 1985) may be described as

follows: A small current ($j_{\rm tot}$) is imposed on a rectangular superconducting film in one direction (the x_1 direction; compare Fig. 13), a small constant temperature gradient (∇T) is imposed across the film in the perpendicular direction (the x_2 direction), and a constant magnetic field $B_{\rm ex}$ is imposed perpendicular to the plane of the film (the x_3 direction). The magnetic flux passing through the film is measured as a function of temperature T. A difference flux (total flux minus total flux for $\nabla T=0$), in the form of a peak (or dip) as a function of temperature, is observed just above the experimentally determined Kosterlitz-Thouless transition. The observed difference flux is linear in the temperature gradient ∇T and the imposed current $j_{\rm tot}$.

The Coulomb gas analogy of this experiment is the following: Vortices correspond to Coulomb gas charges. The current $j_{\rm tot}$ in the x_1 direction corresponds to an electric field E in the x_2 direction acting on the Coulomb gas charges, $E=-j_{\rm tot}4\pi^2\Lambda/(\varphi_0 c)$ [compare Eqs. (4.62) and (6.51)]. The difference in density between positive and negative Coulomb gas charges is given by $\langle \Delta n\rangle=\Delta n_{\rm ex}+\langle \Delta n_g\rangle=B_{\rm ex}/\varphi_0+\Delta B/\varphi_0$ [compare Eq. (4.38)] where $\Delta n_{\rm ex}$ ($\langle \Delta n_g\rangle$) is the charge density associated with the external (the generated difference) magnetic field $B_{\rm ex}$ (ΔB). The temperature gradient in the x_2 direction ∇T is proportional to a temperature gradient in the Coulomb gas temperature $\nabla T^{\rm CG}$ (compare Table II). This gradient gives rise to a Coulomb gas particle-density gradient in the x_2 direction given by

$$\frac{dn}{dx_2}=\frac{dn}{dT^{\rm CG}}\nabla T^{\rm CG} . \tag{7.1}$$

The drift velocity of free positive (negative) Coulomb gas particles $v_{\substack{+\\(-)}}$, $v_+=-v_-$, is given by [compare Eq. (6.1)]

$$v_{\substack{+\\(-)}}(x_2)=\pm\frac{\mathcal{D}}{T^{\rm CG}}E(x_2) . \tag{7.2}$$

The density of free Coulomb gas particles is $n_F(x_2)=n_F^+(x_2)+n_F^-(x_2)$ where the superscript $+$ $(-)$ denotes the density of positive (negative) charges.

The equilibrium condition is the following: The number of positive (negative) particles entering the interval $[x_2,x_2+dx_2]$ per unit time and unit length, $K_{\substack{+\\(-)}}(x_2)$, is given by

$$K_{\substack{+\\(-)}}(x_2)=-\frac{\partial n_F^{\substack{+\\(-)}}}{\partial x_2}v_{\substack{+\\(-)}}(x_2)-n_F^{\substack{+\\(-)}}\frac{\partial v_{\substack{+\\(-)}}}{\partial x_2} . \tag{7.3}$$

In equilibrium one has $K_+(x_2)=K_-(x_2)$. This equilibrium condition follows from the property that Coulomb gas charges can only be created and annihilated in pairs with vanishing separation. This property of the Coulomb gas model is imposed by the vortex analogy because the vorticity in a local region can only change by vortices crossing the boundary of the region (Halperin, 1979).

Equations (7.2) and (7.3), together with the equilibrium condition $K_+(x_2)=K_-(x_2)$, give the balance equation

$$E(x_2)\frac{\partial \ln(n_F)}{\partial x_2}+\frac{T^{\rm CG}}{\mathcal{D}}\frac{\partial}{\partial x_2}\left[\frac{\mathcal{D}}{T^{\rm CG}}E(x_2)\right]=0 . \tag{7.4}$$

The charge density Δn is related to the Maxwell equation $\nabla\cdot\mathbf{E}=2\pi\Delta n$ [Eq. (6.75)] so that

$$\frac{\partial}{\partial x_2}E(x_2)=2\pi\langle \Delta n_g\rangle , \tag{7.5}$$

where $\langle \Delta n_g\rangle=\langle \Delta n\rangle-\Delta n_{\rm ex}$ is the charge density entering the Maxwell equation, because the charge density $\Delta n_{\rm ex}$, corresponding to a vortex density generated by an external magnetic field, is compensated by a uniform background charge. Since the dominant temperature dependence close to the Kosterlitz-Thouless transition is expected to come from n_F, one may safely assume that

$$\frac{\partial}{\partial T^{\rm CG}}\ln(n_F)\gg\frac{\partial}{\partial T^{\rm CG}}\left[\frac{\mathcal{D}}{T^{\rm CG}}\right] .$$

Consequently Eqs. (7.4) and (7.5) can be combined into the approximate condition

$$\langle \Delta n_g\rangle^{\rm av}=-\frac{1}{2\pi}\frac{\partial \ln(n_F)}{\partial T^{\rm CG}}\bigg|_{\rm av}\nabla T^{\rm CG}E_{\rm av} , \tag{7.6}$$

where av denotes spatial average over the x_2 coordinate.

Figure 34 illustrates the qualitative Coulomb gas prediction as obtained by combining Eqs. (2.44) and (7.6). The dominant dependence on the Coulomb gas temperature variable X is given [using a similar type of estimate to Eq. (6.17)] by

$$\Delta n_{\rm ex}+\tilde{C}_1\frac{\partial \ln(n_F)}{\partial X}=n_F[1-\tilde{C}_2(n_F)^{2/X-2}]^{1/2} , \tag{7.7}$$

where $\tilde{C}_1$ and $\tilde{C}_2$ are two constants. Figure 34 shows a typical result for n_F and $\langle \Delta n_g\rangle$ close to the Kosterlitz-

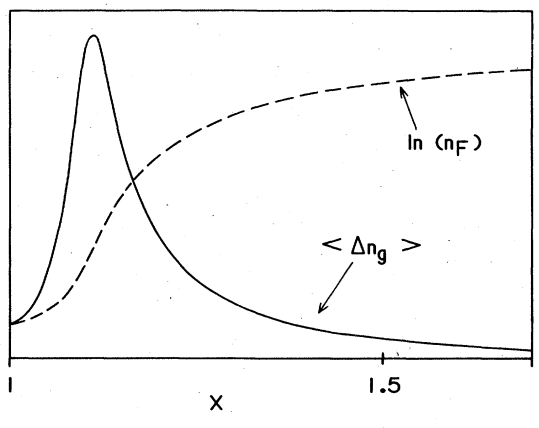

FIG. 34. The generated charge imbalance $\langle \Delta n_g\rangle$ (solid curve) and the logarithm of the free-particle density n_F (dashed curve), as obtained from Eqs. (7.6) and (7.7). The two quantities are plotted in arbitrary units as functions of the temperature variable X. The characteristic feature is the peak of $\langle \Delta n_g\rangle$ just above $X=1$, which arises because $\langle \Delta n_g\rangle$ is proportional to $d\ln(n_F)/dX$.

Thouless transition at $X=1$ in the case $n_{ex} \gg \langle \Delta n_g \rangle$, which corresponds to the typical experimental situation. The characteristic feature is the peak structure of $\langle \Delta n_g \rangle$ just above $X=1$, which arises because $\langle \Delta n_g \rangle$ is proportional to $\partial \ln(n_F)/\partial X$.

The condition given by Eq. (7.6) translated into the superconductor variables becomes [compare Eqs. (4.38), (4.58), (4.62), and (6.51)]

$$\Delta \mathbf{B} = \frac{2\pi\Lambda}{c} \frac{\partial \ln(R/R_N)}{\partial T} \mathbf{j}_{tot} \times \nabla T \ , \tag{7.8}$$

where $\Delta \mathbf{B}$ is the magnetic field difference generated by the temperature gradient ∇T and the current density $\mathbf{j}_{tot}$, R/R_N is the resistance ratio, Λ is the magnetic penetration length, and c the velocity of light. Thus the basic result from the Coulomb gas analogy is that the generated magnetic field difference is, to lowest order, proportional to the product of the current density and the temperature gradient, and the dominant temperature dependence is proportional to the temperature derivative of the logarithm of the resistance ratio. These qualitative results appear to be borne out by the data (Lee *et al.*, 1985; Lee, 1985).

2. Vortex lattice

The vortices generated on a type-II superconducting film by an external perpendicular magnetic field will, for low enough temperatures, form a triangular lattice (Abrikosov, 1957). In the Coulomb gas analogy this corresponds to the formation of a Wigner lattice by two-dimensional Coulomb gas charges of equal sign in a uniform compensating background charge. In case of ^{4}He films the situation corresponds to the formation of a vortex lattice by the vortices associated with a finite rotation velocity, as is expressed by the analogy described in Sec. IV.B.2 [see Eq. (4.30)].

The typical situation for a "dirty" type-II superconducting film is that the vortex lattice is pinned up to the melting temperature T_M, which occurs below the Kosterlitz-Thouless temperature T_c. This means that the melting should be reflected in a sudden increase of dissipation, since above the melting temperature T_M most of the vortices are free to move. In the experiment by Hebard and Fiory (1980), which, as described in Sec. VI.B.2, measures the Coulomb gas dielectric constant $\varepsilon(\omega)$, the dissipation is proportional to the imaginary part of $\varepsilon(\omega)$. A sudden increase of the imaginary part of $\varepsilon(\omega)$, in the limit of small ω, was observed by Fiory and Hebard (1982) at a temperature below the experimentally determined Kosterlitz-Thouless temperature and was interpreted as being caused by the melting of the vortex lattice in the way described above. Similar observations have been made by Martinoli *et al.* (1982), who gave the same qualitative explanation.

The melting of the two-dimensional vortex lattice has been considered by Huberman and Doniach (1979) and in more detail by Fisher (1980). Its basic features are the following: The Wigner lattice for two-dimensional Coulomb gas charges is incompressible (due to the logarithmic particle interaction) and the shear modulus at zero temperature μ^S is given by (the Coulomb gas charge is unity in our unit system)

$$\mu^S = \langle \Delta n \rangle / 8 \ , \tag{7.9}$$

where $\langle \Delta n \rangle$ is the charge density. For a superconducting film this corresponds to the limiting case $\Lambda^{-2} \ll B_{ex}/\varphi_0 \ll \xi^{-2}$ (Conen and Schmid, 1974). The area of a single triangle is $(\sqrt{3}/4)a^2$ in the triangular lattice where a is the lattice constant, so that $\langle \Delta n \rangle = 4/(\sqrt{3}a^2)$ and consequently

$$\mu^S = (2\sqrt{3}a^2)^{-1} \ . \tag{7.10}$$

The energy of a single dislocation, corresponding to a Burger's vector of one lattice spacing, is in the continuum elasticity theory (see, for example, Kosterlitz and Thouless, 1973) given by

$$E_{dis} = \frac{\mu^S a^2}{8\pi} \ln(\Omega/a^2) \ , \tag{7.11}$$

where Ω is the total area of the Wigner lattice. The free energy for a single dislocation is, in analogy with Eq. (2.23), given by (Kosterlitz and Thouless, 1973)

$$F_{dis} = \left[\frac{\mu^S a^2}{8\pi} - T^{CG} \right] \ln(\Omega/a^2) \ . \tag{7.12}$$

In analogy with Eq. (2.23), the shear modulus in the presence of dislocations, μ_R^S, may then be expected to vanish due to thermally created free dislocations for $T > T_M$, where

$$T_M^{CG} = \frac{\mu^S a^2}{8\pi} = \frac{1}{16\pi\sqrt{3}} \ . \tag{7.13}$$

This estimate of the melting temperature for a two-dimensional lattice is due to Kosterlitz and Thouless (1973) and the specialization to the Wigner lattice for two-dimensional Coulomb gas charges to Huberman and Doniach (1979) and Fisher (1980).

The description of the thermally created dislocations turns out to be analogous to that of vortices, and a possible mechanism for the melting is unbinding of dislocation pairs (Kosterlitz and Thouless, 1973).

Provided this is the mechanism of the melting, then the estimate of the melting temperature given in Eq. (7.13) can be improved in two ways: First, the small-wavelength fluctuations should be included, which changes $\mu^S = \mu^S(0)$ to $\mu^S(T_M^{CG}) < \mu^S(0)$ (Fisher, 1980). Second, the polarization due to dislocation pairs has to be included; it can be described by a dielectric constant $\tilde{\varepsilon}_{dis} > 1$ analogous to $\tilde{\varepsilon}$ describing polarization due to vortex pairs. The relation between the Kosterlitz-Thouless temperature T_c^{CG} and the melting temperature T_M^{CG} can hence be estimated as [compare Eqs. (2.23) and (7.13)]

$$T_M^{CG}=AT_c^{CG}\ , \tag{7.14a}$$

with the constant A given by

$$A=\frac{1}{4\pi\sqrt{3}}\frac{\mu^S(T_M^{CG})}{\mu^S(0)}\frac{\tilde{\varepsilon}(T_c^{CG})}{\tilde{\varepsilon}_{dis}(T_M^{CG})}\ , \tag{7.14b}$$

where $\tilde{\varepsilon}$ is the Coulomb gas dielectric constant and $\tilde{\varepsilon}_{dis}$ is the corresponding constant for dislocations. Fisher (1980) obtained the estimate $0.4<\mu^S(T_M^{CG})/[\mu^g(0)\tilde{\varepsilon}_{dis}(T_M^{CG})]<0.75$. An estimate of the real melting temperature T_M can then be obtained from the approximate temperature dependence given by Eq. (4.59),

$$T_M=\frac{T_cT_{c0}A}{T_{c0}-(1-A)T_c}\ . \tag{7.15}$$

If the melting is due to dislocation unbinding, then, according to the two-dimensional melting theory by Halperin, Nelson, and Young (Halpern and Nelson, 1978; Nelson and Halperin, 1979; Young, 1979), the melting is into a hexatic phase with long-range angular correlation, which subsequently melts into a disordered phase by an analogous disclination unbinding transition. On the other hand, substrate interactions like vortex pinning can change the transition into a first-order transition, in which case the estimate of the melting temperature T_M given by Eq. (7.15) turns into an upper bound.

The actual situation for the vortex lattice in the case of "dirty" type-II superconducting films appears to be the following: The melting occurs at a temperature T_M consistent with Eq. (7.15) (Fiory and Hebard, 1982; Hebard and Fiory, 1982; Martinoli *et al.*, 1982). Pinning effects on the melting transition appear to be pronounced (Martinoli *et al.*, 1982).

3. Connection to BCS theory

As described in Sec. II.B.1, the Kosterlitz-Thouless temperature for the two-dimensional Coulomb gas is given by

$$T_c^{CG}=\frac{1}{4\varepsilon_c}\ .$$

The Coulomb gas temperature is related to the real temperature by

$$T^{CG}=T\Big/\left[2\pi\rho\left[\frac{\hbar}{m^*}\right]^2\right]$$

[see Eq. (4.20) and Table II]. The perpendicular magnetic penetration depth Λ is, within a Ginzburg-Landau description, related to the superfluid mass density [see Eq. (4.32b)] by

$$\Lambda=\frac{1}{2\pi}\left[\frac{m^*c}{e^*}\right]^2\frac{1}{\rho}\ .$$

Provided that the BCS theory is applicable, the perpendicular penetration depth Λ can also be expressed in terms of the BCS energy gap $\Delta(T)$. For a "dirty" type-II superconducting film, this relation can be expressed (see, for example, Tinkham, 1975) as

$$\Lambda=\frac{\varphi_0^2}{2\pi^4}\frac{R_Ne^2}{\hbar}\left[\Delta(T)\tanh\left[\frac{\Delta(T)}{2T}\right]\right]^{-1}\ , \tag{7.16}$$

where R_N is the normal-state sheet resistance of the film. Combining Eq. (7.16) with the relations quoted so far in this section gives

$$\frac{R_N}{R_c}=\frac{\pi^2\Delta(T)\tanh[\Delta(T)/2T]}{8T_c\varepsilon_c}\simeq\frac{5.78}{\varepsilon_c}\left[\frac{T_{c0}}{T_c}-1\right]\ , \tag{7.17}$$

where $R_c=\hbar/e^2$ corresponds to a sheet resistance of $R_c=4.11$ $(k\Omega/m^2)$ and the last approximate equality is obtained from the relation $\Delta(T)\simeq3.06T_{c0}(1-T/T_{c0})^{1/2}$ valid close to the BCS temperature T_{c0}. Thus this chain of relations leads to a simple approximate relation between the Kosterlitz-Thouless temperature T_c, the mean-field temperature T_{c0}, and the normal-state sheet resistance R_N given by (Beasley *et al.*, 1979; Lozovik and Akapov, 1981; Epstein *et al.*, 1982)

$$\frac{T_c}{T_{c0}}=\left[1+0.173\varepsilon_c\frac{R_N}{R_c}\right]^{-1}\ . \tag{7.18}$$

This relation is in qualitative agreement with the experimental result for "dirty" type-II superconducting films (Beasley *et al.*, 1979; Epstein *et al.*, 1982; Kadin *et al.*, 1983).

VIII. FINAL REMARKS

The scope of the present article has been the two-dimensional Coulomb gas and its connection to vortex fluctuations for superfluid-superconducting films. The subject has been limited to the case when the superfluid on some length scale can be described in terms of a phenomenological Ginzburg-Landau-type theory with vortices (see Sec. IV.D). This means that all the physics associated with smaller length scales is hidden in the phenomenological parameters.

For example, coupled arrays of Josephson junctions and granular superconducting films involve interesting problems associated with the array structure, which are not discussed in the present article. For a review of such problems see, for example, Doniach (1984), Lobb (1984), and Mooij (1984). This area of study may involve questions about the relation between the properties of a single junction and the phenomenological parameters entering on a Ginzburg-Landau level of description, or the effects of disorder in the case of granular superconducting films. One interesting aspect of the subject is the effect of zero-point-charge fluctuations, across the junctions or between the superconducting grains, on the Kosterlitz-Thouless

transition and the possibility of a reentrant behavior [for a review, see, for example, Doniach, 1984; more recently Jacobs *et al.* (1984) suggest a first-order reentrant behavior].

Another aspect of this topic is the interplay between the array structure and the vortices imposed by a perpendicular external magnetic field (for a review, see Lobb, 1984). This forms an entire subject by itself outside the scope of the present article.

The relation to the phenomenological description reviewed in the present article is that on some length scale ξ, larger than the scale of the array structure, one may expect that the phenomenological description applies as long as $B_{ex} \ll \varphi_0/\xi^2$ (compare Sec. IV.E).

Considerable progress has been made in developing a theory of vortex fluctuations for superconducting films over the last ten years, and the present stage is that the two-dimensional Coulomb gas provides a convincing qualitative explanation of many experimental observations for superconducting films and ^{4}He films.

Nevertheless many open questions remain, and from this point of view the subject is still in some respects at an early stage. For example, the properties of the two-dimensional Coulomb gas in itself are not as yet known with great precision, except for the critical properties precisely at the Kosterlitz-Thouless transition. Neither is the detailed connection yet known between the microscopic level of description for ^{4}He films, superconducting films, and two-dimensional arrays of Josephson junctions, on the one hand, and the Ginzburg-Landau level of description (which forms a convenient starting point for the Coulomb gas analogy) on the other.

Much experimental work so far has been designed to test directly the critical properties at the Kosterlitz-Thouless transition. However, as discussed, such tests may not be as conclusive as one would like for the following reasons: Experiments are by necessity performed on finite samples, which strictly speaking have no true critical properties. From a practical point of view this means that the critical properties for temperatures too *close* to the critical temperature are masked by finite length scales, either of the sample size or, as may be the case for superconducting films, of the perpendicular magnetic penetration length Λ. On the other hand, for temperatures too *far* from the critical temperature one is outside the critical region. Hence the possibility that the critical behavior derived from the renormalization-group equations applies only to a very small temperature window has to be considered. Second, the explicit functional forms obtained from the renormalization-group equations usually involve many free parameters in relation to the structure of the data, so that the significance even of a very good fit to the data is somewhat uncertain.

Consequently it is probably fair to state that the functional forms derived from the renormalization-group equations provide a qualitative, but not necessarily a true quantitative, explanation of the experiments so far.

On a more phenomenological level, the Coulomb gas scaling relations have provided an alternative way of establishing the connection between experiments and theory. These scaling relations seem to be well obeyed by experiments. However, the degree of validity of these various scaling relations has not yet been carefully investigated through systematic experiments. For example, one such possibility of a systematic investigation for superconductors would be to perform a series of experiments on samples that, outside a small temperature interval around the Kosterlitz-Thouless temperature, are manifestly well described by standard Ginzburg-Landau theory.

In summary, vortex unbinding is reflected in many experiments on superconducting films, two-dimensional arrays of Josephson junctions, and ^{4}He films. However, the extent to which the theoretical description, in terms of the two-dimensional Coulomb gas and the renormalization-group equations for the Kosterlitz-Thouless transition, provides a true quantitative explanation of the experiments is as yet largely unknown. In view of this, the subject may be expected to remain an experimental and theoretical challenge for some time to come.

REFERENCES

Abraham, D. W., C. J. Lobb, M. Tinkham, and T. M. Klapwijk, 1982, Phys. Rev. B **26**, 5268.

Abrikosov, A. A., 1957, Zh. Eksp. Teor. Fiz. **32**, 1442 [Sov. Phys. JETP **5**, 1174 (1957)].

Adams, P. W., and W. I. Glaberson, 1986, Phys. Rev. Lett. **57**, 82.

Agnolet, G., 1983, Ph.D. thesis (Cornell University).

Agnolet, G., D. C. McQueeney, and J. D. Reppy, 1984a, in *Proceedings of the 17th International Conference on Low Temperature Physics,* edited by U. Eckern, A. Schmid, W. Weber, and H. Wuhl (North-Holland, Amsterdam), Pt. II, p. 965.

Agnolet, G., D. C. McQueeney, and J. D. Reppy, 1984b, unpublished.

Agnolet, G., S. Teitel, and J. D. Reppy, 1981, Phys. Rev. Lett. **47**, 1537.

Ambegaokar, V., B. I. Halperin, D. R. Nelson, and E. D. Siggia, 1978, Phys. Rev. Lett. **40**, 783.

Ambegaokar, V., B. I. Halperin, D. R. Nelson, and E. D. Siggia, 1980, Phys. Rev. B **21**, 1806.

Ambegaokar, V., and S. Teitel, 1979, Phys. Rev. B **19**, 1667.

Ambit, D. J., 1984, *Field Theory, the Renormalization Group, and Critical Phenomena,* 2nd ed. (World Scientific, Singapore).

Amit, D. J., Y. Y. Goldschmidt, and G. Grinstein, 1980, J. Phys. A **13**, 585.

Aslamazov, L. G., and A. I. Larkin, 1968, Phys. Lett. A **26**, 238.

Bancel, P. A., and K. E. Gray, 1981, Phys. Rev. Lett. **46**, 148.

Bardeen, J., and M. J. Stephen, 1965, Phys. Rev. **140**, 1197A.

Baym, G., 1969, in *Mathematical Models of Solid State and Superfluid Theory,* edited by R. C. Clark (Oliver and Boyd, Edinburgh), p. 139.

Beasley, M. R., J. E. Mooij, and T. P. Orlando, 1979, Phys. Rev. Lett. **42**, 1165.

Berezinskii, V. L., 1971, Zh. Eksp. Teor, Fiz. **61**, 1144 [Sov. Phys. JETP **34**, 610 (1972)].

Bergman, D., 1969, Phys. Rev. **188**, 370.

Bergman, D., 1971, Phys. Rev. A **3**, 2058.

Berker, A. N., and D. R. Nelson, 1979, Phys. Rev. B **19**, 2488.

Berthold, J. E., D. J. Bishop, and J. D. Reppy, 1977, Phys. Rev. Lett. **39**, 348.
Bishop, D. J., J. E. Berthold, J. M. Parpia, and J. D. Reppy, 1981, Phys. Rev. B **24**, 5047.
Bishop, D. J., and J. D. Reppy, 1978, Phys. Rev. Lett. **40**, 1727
Bishop, D. J., and J. D. Reppy, 1980, Phys. Rev. B **22**, 5171.
Browne, D. A., and S. Doniach, 1982, Phys. Rev. B **25**, 136.
Caillol, J. M., and D. Levesque, 1986, Phys. Rev. B **33**, 499.
Chui, S. T., and P. A. Lee, 1975, Phys. Rev. Lett. **35**, 315.
Clem, J. R., 1979, in *Inhomogeneous Superconductors—1979,* AIP Conference Proceedings No. 58, edited by D. U. Gubser, T. Francavilla, S. A. Wolf, and J. R. Leibowitz (AIP, New York), p. 245.
Coleman, S., 1975, Phys. Rev. D **11**, 2088.
Conen, E., and A. Schmid, 1974, J. Low Temp. Phys. **17**, 331.
Côté, R., and A. Griffin, 1986, Phys. Rev. B **34**, 6240.
Crooker, B. C., B. Hebral, E. N. Smith, Y. Takano, and J. D. Reppy, 1983, Phys. Rev. Lett. **51**, 666.
Deutsch, C., and M. Lavaud, 1974, Phys. Rev. A **9**, 2598.
Doniach, S., 1984, in *NATO Advanced Study Institute on Localization, Percolation, and Superconductivity,* edited by A. M. Goldman and S. A. Wolf (Plenum, New York), p. 401.
Doniach, S. and B. A. Hubermann, 1979, Phys. Rev. Lett. **42**, 1169.
Epstein, K., A. M. Goldman, and A. M. Kadin, 1981, Phys. Rev. Lett. **47**, 534.
Epstein, K., A. M. Goldman, and A. M. Kadin, 1982, Phys. Rev. B **26**, 3950.
Everts, H. U., and W. Koch, 1977, Z. Phys. B **28**, 117.
Fetter, A. L., 1980, Phys. Rev. B **22**, 1200.
Finotello, D., and F. M. Gasparini, 1985, Phys. Rev. Lett. **55**, 2156.
Fiory, A. T., and A. F. Hebard, 1982, Phys. Rev. B **25**, 2073.
Fiory, A. T., A. F. Hebard, and W. I. Glaberson, 1983, Phys. Rev. B **28**, 5075.
Fisher, D. S., 1980, Phys. Rev. B **22**, 1190.
Fisher, M. E., M. N. Barber, and D. Jasnow, 1973, Phys. Rev. A **8**, 1111.
Forster, D., 1975, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Benjamin, Reading, MA).
Friedman, H. L., 1962, *Ionic Solution Theory* (Interscience, New York).
Frölich, J., 1976, in *Renormalization Theory: Proceedings of the NATO Advanced Study Institute held at the International School of Mathematical Physics. . .Erice. . .1975,* edited by G. Velo and A. S. Wightman (Reidel, Dordrecht/Boston), p. 371.
Garland, J. C., and D. J. VanHarlingen, Phys. Rev. Lett. **55**, 2047.
Glaberson, W. I., and R. J. Donelly, 1986, in *Progress in Low Temperature Physics Vol. IX,* edited by D. F. Brewer (North-Holland, Amsterdam), p. 5.
Gray, K. E., 1986, Phys. Rev. Lett. **56**, 2879.
Gray, K. E., J. Brorson, and P. A. Bancel, 1985, J. Low Temp. Phys. **59**, 329.
Grinstein, G., P. Minnhagen, and A. Rosengren, 1979, J. Phys. C **12**, 1271.
Halperin, B. I., 1979, in *Physics of Low-Dimensional Systems,* Proceedings of Kyoto Summer Institute, Sept. 1979, edited by Y. Nagaoka and S. Hikami (Publication Office, Prog. Theor. Phys., Kyoto), p. 53.
Halperin, B. I., and D. R. Nelson, 1978, Phys. Rev. Lett. **41**, 121.
Halperin, B. I., and D. R. Nelson, 1979, J. Low Temp. Phys. **36**, 599.
Hauge, E. H., and P. C. Hemmer, 1971, Physica Norvegica **5**, 209.
Hebard, A. F., and A. T. Fiory, 1980, Phys. Rev. Lett. **44**, 291.
Hebard, A. F., and A. T. Fiory, 1982, in *Proceedings of the 16th International Conference on Low Temperture Physics,* edited by W. G. Clark (Physica **109&110B+C**, 1637).
Hebard, A. F., and A. T. Fiory, 1983, Phys. Rev. Lett. **50**, 1603.
Hohenberg, P. C., 1971, in *Proceedings of the International School of Physics Enrico Fermi, Course 51,* edited by M. S. Green (Academic, New York/London), p. 285.
Hohenberg, P. C., and P. C. Martin, 1965, Ann. Phys. **34**, 291.
Hu, C.-R., 1972, Phys. Rev. B **6**, 1756.
Huberman, B. A., and S. Doniach, 1979, Phys. Rev. Lett. **43**, 950.
Huebner, R. P., 1979, *Magnetic Flux Structures in Superconductors* (Springer, Berlin).
Høye, J. S., and K. Olaussen, 1980, Physica A **104**, 447.
Høye, J. S., and K. Olaussen, 1981, Physica A **107**, 241.
Jacobs, L., J. V. José, and M. A. Novotny, 1984, Phys. Rev. Lett. **53**, 2177.
José, J. V., L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson, 1977, Phys. Rev. B **16**, 1217.
Joseph, R. A., and F. M. Gasparini, 1982, Phys. Rev. B **26**, 3982.
Josephson, B. D., 1962, Phys. Lett. **1**, 251.
Kadin, A. M., K. Epstein, and A. M. Goldman, 1983, Phys. Rev. B **27**, 6691.
Kihmi, D., F. Leyvraz, and D. Ariosa, 1984, Phys. Rev. B **29**, 1487.
Kim, Y. B., and M. J. Stephen, 1969, in *Superconductivity 2,* edited by R. D. Parks (Dekker, New York), p. 1107.
Kosterlitz, J. M., 1974, J. Phys. C **7**, 1046.
Kosterlitz, J. M., and D. J. Thouless, 1972, J. Phys. C **5**, L124.
Kosterlitz, J. M., and D. J. Thouless, 1973, J. Phys. C **6**, 1181.
Kosterlitz, J. M., and D. J. Thouless, 1978, in Prog. Low Temp. Phys. B **7**, 371.
Kotsubo, V., and G. A. Williams, 1984, Phys. Rev. Lett. **53**, 691.
Kotsubo, V., and G. A. Williams, 1986, Phys. Rev. B **33**, 6106.
Lee, H. J., 1985, Ph.D. thesis (Ohio State University).
Lee, H. J., D. A. Rudman, and J. C. Garland, 1985, Phys. Rev. Lett. **55**, 2051.
Lifshitz, E. M., and L. P. Pitaevskii, 1980, *Statistical Physics Part 2,* 2nd ed. (Pergamon, New York).
Lobb, C. J., 1984, in *Proceedings of the 17th International Conference on Low Temperature Physics,* edited by U. Eckern, A. Scmid, W. Weber, and H. Wuhl (Physica **126B+C**, 319).
Lozovik, Y. E., and S. G. Akapov, 1980, Solid State Commun. **35**, 693.
Lozovik, Y. E., and S. G. Akapov, 1981, J. Phys. C **14**, L31.
Luther, A. H., and V. J. Emery, 1974, Phys. Rev. Lett. **33**, 589.
Luther, A. H., and I. Peschel, 1974, Phys. Rev. B **9**, 2911.
Maps, J., and R. B. Hallock, 1981, Phys. Rev. Lett. **47**, 1533.
Maps, J., and R. B. Hallock, 1982, Phys. Rev. B **26**, 3979.
Martinoli, P., M. Nsabimana, G.-A. Racine, H. Beck, and J. R. Clem, 1982, Helv. Phys. Acta **55**, 655.
Masker, W. E., S. Marcelja, and R. D. Parks, 1969, Phys. Rev. **188**, 745.
Matsubara, T., and H. Matsuba, 1956, Prog. Theor. Phys. **16**, 416.
Matsubara, T., and H. Matsuba, 1957, Prog. Theor. Phys. **17**, 19.
McQueeney, D. C., G. Agnolet, and J. D. Reppy, 1984, Phys. Rev. Lett. **52**, 1325.

Menyhard, N., and J. Solyom, 1973, J. Low Temp. Phys. **12**, 529.
Mermin, N. D., and H. Wagner, 1966, Phys. Rev. Lett. **17**, 1133.
Minnhagen, P., 1980, Solid State Commun. **36**, 805.
Minnhagen, P., 1981a, Phys. Rev. B **23**, 5745.
Minnhagen, P., 1981b, Phys. Rev. B **24**, 6758.
Minnhagen, P., 1983a, Phys. Rev. B **27**, 2807.
Minnhagen, P., 1983b, Phys. Rev. B **28**, 2463.
Minnhagen, P., 1984a, in *NATO Advanced Study Institute on Localization, Percolation, and Superconductivity,* edited by A. M. Goldman and S. A. Wolf (Plenum, New York), p. 287.
Minnhagen, P., 1984b, Phys. Rev. B **29**, 1440.
Minnhagen, P., 1985a, Phys. Rev. Lett. **54**, 2351.
Minnhagen, P., 1985b, Phys. Rev. B **32**, 3088.
Minnhagen, P., 1985c, unpublished.
Minnhagen, P., 1986, unpublished.
Minnhagen, P., and M. Nylén, 1985a, Phys. Rev. B **31**, 5768.
Minnhagen, P., and M. Nylén, 1985b, Phys. Rev. B **31**, 1693.
Minnhagen, P., A. Rosengren, and G. Grinstein, 1978, Phys. Rev. B **18**, 1356.
Minnhagen, P., and M. Wallin, 1987, Phys. Rev. B (in press).
Minnhagen, P., and G. G. Warren, 1981, Phys. Rev. B **24**, 2526.
Minnhagen, P., and H. Weber, 1985, Phys. Rev. B **32**, 3337.
Mooij, J. E., 1983, in *NATO Advanced Study Institute on Advances in Superconductivity,* edited by B. Deaver and J. Ruvalds (Plenum, New York), p. 433.
Mooij, J. E., 1984, in *NATO Advanced Study Institute on Percolation, Localization, and Superconductivity,* edited by A. M. Goldman and S. A. Wolf (Plenum, New York), p. 325.
Myerson, R. J., 1978, Phys. Rev. B **18**, 3204.
Nelson, D. R., 1978, Phys. Rev. B **18**, 2318.
Nelson, D. R., 1980, in *Fundamental Problems in Statistical Mechanics V,* edited by E. G. D. Cohen (North-Holland, New York), p. 53.
Nelson, D. R., 1983, in *Phase Transitions and Critical Phenomena 7,* edited by C. Domb and J. L. Lebowitz (Academic, London), p. 1.
Nelson, D. R., and B. I. Halperin, 1979, Phys. Rev. B **19**, 2457.
Nelson, D. R., and J. M. Kosterlitz, 1977, Phys. Rev. Lett. **39**, 1201.
Nozières, P., and W. F. Vinen, 1966, Philos. Mag. **14**, 667.
Ohta, T., 1978, Prog. Theor. Phys. **60**, 968.
Ohta, T., and D. Jasnow, 1979, Phys. Rev. B **20**, 139.
Pearl, J., 1964, Appl. Phys. Lett. **5**, 65.
Pearl, J., 1965, in *Low Temperature Physics—LT9,* edited by J. G. Daunt, D. O. Edwards, F. J. Milford, and M. Yagub (Plenum, New York), p. 566.
Polyakov, A. M., 1977, Nucl. Phys. B **20**, 429.
Rasolt, M., M. J. Stephen, M. E. Fisher, and P. B. Weichman, 1984, Phys. Rev. Lett. **53**, 798.
Reppy, J. D., 1984, in *Proceedings of the 17th International Conference on Low Temperature Physics,* edited by U. Eckern, A. Scmid, W. Weber, and H. Wuhl (Physica **126B+C**, 335).
Resnick, D. J., J. C. Garland, J. T. Boyd, S. Shoemaker, and R. S. Newrock, 1981, Phys. Rev. Lett. **47**, 1542.
Rudnick, I., 1978, Phys. Rev. Lett. **40**, 1454.
Samuel, S., 1978, Phys. Rev. D **18**, 1916.
Savit, R., 1978, Phys. Rev. B **17**, 1340.
Shenoy, S. R., 1985a, J. Phys. C **18**, 5143.
Shenoy, S. R., 1985b, J. Phys. C **18**, 5163.
Smith, E. N., D. J. Bishop, J. E. Berthold, and J. D. Reppy, 1978, J. Phys. (Paris) Colloq. **C6**, 342.
Suzuki, M., 1979, in *Physics of Low-Dimensional Systems,* Proceedings of Kyoto Summer Institute, edited by Y. Nagaoka, and S. Hikami (Publication Office, Prog. Theor. Phys., Kyoto), p. 39.
Svistunov, V. M., A. I. D'Jachenko, and V. Y. Tarenkov, 1984, J. Low Temp. Phys. **57**, 619.
Teitel, S., 1981, Ph.D. thesis (Cornell University).
Teitel, S., 1982, J. Low Temp. Phys. **46**, 77.
Teitel, S., and C. Jayaprakash, 1983, Phys. Rev. B **27**, 598.
Tinkham, M., 1975, *Introduction to Superconductivity* (McGraw-Hill, New York).
Trugman, S. A., and S. Doniach, 1982, Phys. Rev. B **26**, 3682.
Turkevich, L. A., 1979, J. Phys. C **12**, L385.
van Himbergen, J. E., and S. Chakravarty, 1981, Phys. Rev. B **23**, 359.
Vinen, W. F., 1961, Prog. Low Temp. Phys. **3**, 1.
Vinen, W. F., 1969, in *Superconductivity Vol. II,* edited by R. D. Parks (Dekker, New York), p. 1167.
Voss, R. F., and R. A. Webb, 1982, Phys. Rev. B **25**, 3446.
Wang, C., and L. Yu, 1986, Phys. Rev. B **33**, 599.
Weichman, P. B., M. Rasolt, M. E. Fisher, and M. J. Stephen, 1986, Phys. Rev. B **33**, 4632.
Wiegman, P. B., 1978, J. Phys. C **11**, 1583.
Wigner, E., 1938, Trans. Faraday Soc. **34**, 678.
Wilks, J., 1967, *The Properties of Liquid and Solid Helium* (Clarendon, Oxford), p. 335.
Wilson, K. G., 1971, Phys. Rev. B **4**, 3174.
Wolf. S. A., D. U. Gubser, W. W. Fuller, J. C. Garland, and R. S. Newrock, 1981, Phys. Rev. Lett. **47**, 1071.
Young, A. P., 1978, J. Phys. C **11**, L453.
Young, A. P., 1979, Phys. Rev. B **19**, 1855.
Young, A. P., 1980, in *NATO Advanced Study Institute on Ordering in Strongly Fluctuating Condensed Matter Systems* (Plenum, New York), p. 271.
Young, A. P., and T. Bohr, 1981, J. Phys. C **14**, 2713.
Yu, L., 1982, Phys. Rev. B **25**, 198.
Zittartz, J., 1976, Z. Phys. B **23**, 277.